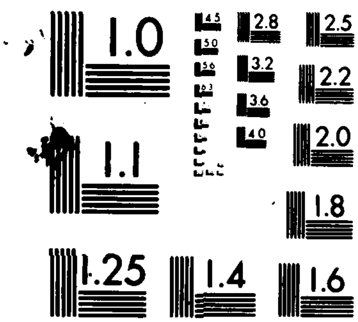
AD-A113 374 CONTROL DATA CORP MONTEREY CA F/G 4/2
TIME SCALES OF SYSTEMATIC ERRORS IN A NUMERICAL WEATHER FORECAS--ETC(U)
JAN 82 P A HARR N00228-80-C-GE83
UNCLASSIFIED NEPRF-CR-82-03 NL

1 - 1

END
DATE
FILMED
4 82
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD A113374

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NAVENVPREDRSCHFAC Contractor Report CR 82-03	2. GOVT ACCESSION NO. AD-A113 374	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Time Scales of Systematic Errors in a Numerical Weather Forecast		5. TYPE OF REPORT & PERIOD COVERED Final
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Patrick A. Harr		8. CONTRACT OR GRANT NUMBER(s) N00228-80-C-QE83
9. PERFORMING ORGANIZATION NAME AND ADDRESS Control Data Corporation 205 Montecito Avenue Monterey, CA 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS PE 63207N PN 7W0513 TA CCOO NEPRF WU 6.3-13
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Air Systems Command Department of the Navy Washington, DC 20361		12. REPORT DATE January 1982
		13. NUMBER OF PAGES 58
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Naval Environmental Prediction Research Facility Monterey, CA 93940		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Forecast errors
Random errors
Time scale

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Many numerical model verification systems are handicapped by their inability to separate random errors, defined as scales of motion smaller than the concerned scale, and systematic errors. Random errors may be eliminated by time averaging over a minimum allowable data period.

Statistical methods are implemented to determine a minimum data period needed to constitute a time scale within which numerical forecast errors are truly systematic and not smoothed fields of rapidly varying random errors for a synoptic scale event. (Continued on reverse.)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Block 20, Abstract, continued.

Long term or model climatological error patterns of a specific event are compared with randomly determined shorter term error patterns through statistical tests and measures to provide results which indicate a minimum case number needed to sufficiently smooth out unwanted random errors.

TABLE OF CONTENTS

1.0 Introduction

Numerical models provide most of the foundation of synoptic scale weather predictions produced by large weather forecasting and data centers. It is therefore evident that the quality and accuracy of environmental forecasts are dependent upon the quality and accuracy of the numerical simulations of the atmosphere. Many studies have been conducted for the purpose of examining the accuracy of numerical model simulations. The great majority of model verification studies are aimed at the synoptic scale event and usually single out a specific case study of some weather situation. Also most verification studies involve two sets of examination. One set is usually based on statistical measures such as a root mean square error (RMSE), correlation or other commonly used scores (eg. threat score, S1 score, etc.) The second set can involve more complex analysis such as spectrum analysis, diagnostic studies and other tools which are generallv more complex and costly than simple statistical measures.

There are two types of errors which may occur in a forecasting system. These are systematic errors and random errors. Systematic errors can be examined and analyzed with respect to their sources. Examples would be errors due to grid spacing or errors due to the finite differencing formulations. For instance, finite difference approximations to the differential equations used to simulate the atmosphere can cause phase speed errors of shorter wavelength waves in the atmospheric simulations. Random errors, however, cannot be examined in relation to their sources. We can define random errors as being associated with the stochastic nature of the atmosphere. We can also classify random errors as errors associated with scales of motion smaller than the resolvable scales the study or model is concerned with.

Most verification methodologies or schemes commonly utilized for studies of model accuracy are hindered by their inability to distinguish between the systematic errors and random errors within the simulated data field. It is therefore necessary to try and eliminate the random errors. Based on the above classification of random errors, they may be eliminated by time averaging over a certain data period. However the identification of a data period which sufficiently smoothes random errors but not systematic errors is not easy.

This report documents a methodology, which is based on statistical methods and can identify the data period which identifies the time scales over which systematic errors occur but random errors are sufficiently smoothed.

The resultant time scales shall define error patterns which are truly systematic and not smoothed fields of rapidly varying random errors.

1.1 Objectives

The primary objective of this study is to define a data period which may be averaged to sufficiently smooth random errors in a numerical forecast.

Certain characteristic error patterns between forecasts and verifying analysis of specific synoptic features which are based upon various time scales are produced. Error characteristics are highlighted for the longer term, climatological period, and shorter time periods obtained by randomly grouping forecasts and analyses respectively. Three resultant cases are possible.

If error characteristics, defined as the spatial variation in differences between forecast and verifying analysis values and the amplitude of these values, are uniform over all time scales (e.g. climatological and shorter periods) we conclude that the errors are truly systematic and not smoothed random errors.

Another possibility may be that error characteristics are uniform for all but, say, the shorter time scales of 1-3 days. This indicates that it may be necessary to average data for periods longer than 3 days (or cases) in order to insure that the random error components are sufficiently smoothed, enabling an investigation of the systematic error components using other widely used verification techniques.

The final case would be where error characteristics for all time scales are not uniform. This would indicate that the error associated with the particular scale is not systematic but rather random and could not be examined using conventional verification techniques.

This report defines the statistical procedure which may be used to define the appropriate time scale. Results of test applications of the procedure are presented as an aid in highlighting the specific details of the methodology. Data was gathered from the Fleet Numerical Oceanography Center's (FNOC) Primitive Equation Model. No significant attempt is made to interpret the results in terms of this particular model.

2.0 Time Scale Methodology

The statistical methodology is aimed at defining error patterns for different time scales. The time scales shall range from the long term or "climatological" scale down to an average of 3 fields. Comparisons between error patterns of different groups and the actual error patterns are based upon objective statistical measures.

2.1 Data

This study deals with synoptic scale weather events. For this purpose and for test applications the Shanghai Low synoptic event was chosen. This was chosen because it is a well observed case of cyclogenesis plus there have been well documented errors which pertain to the synoptic situation. A general application of this technique could utilize data pertaining to general synoptic scale situations.

Grid point data was obtained from the climatological data base at FNOC. The 24, 48 and 72 hour 1000 MB forecasts and verifying analyses were obtained for 30 identified cases of the Shanghai Low. The cases were chosen from the time periods January through March of 1978 and December, 1979 through March 1980. The well defined and sometimes explosive cyclogenesis common to the China coast during the winter months provide a well suited data set for an application of this type of methodology.

Shanghai Lows were identified as closed systems which formed and intensified between 110E and 150E longitude and 20N and 50N latitude. All cases were initially chosen on either a 00Z or 12Z time boundary.

Data fields were chosen for the following quantities and time periods for any specific case:

Time	Fields Chosen
time of formation	current analysis 24,48,72 hour forecasts
12 hours later	current analysis 24,48,72 hour forecasts
24 hours later	current analysis 24,48 hour forecasts
36 hours later	current analysis
48 hours later	current analysis
60 hours later	current analysis
72 hours later	current analysis
84 hours later	current analysis

For the 24, 48 and 72 hour forecasts two time periods are available for each case. These are identified as the initial 12 hour period and the second 12 hour period. Classification of the data in this manner may give a better sequence of error characteristics associated with the initial cyclogenesis process.

2.2 Analysis

The data set is analyzed for each category, consisting of a forecast interval(eg. 24,48 or 72 hrs.) and time period(eg. initial 12 hrs. or second 12 hrs.) respectively.

The initial analysis involves the averaging of all cases contained in each category defined above and thereby producing an averaged forecast and analysis field composed of a unique number of cases for each time period respectively. We define these fields as the "climatological" forecast and analysis for our system. The 24, 48 and 72 hour "climatological" forecast and analysis, likewise defines the "climatological" error for the chosen synoptic event.

The error patterns for the climatological group can be identified by simply differencing the resulting averaged forecast and verifying analysis fields. However a single difference field provides no information as to the significance of the error. If the chosen synoptic event or data region includes an area which has a high natural variance, the difference field will be "biased" by the domination of the variable area. For example if a chosen analysis region should include the area around the mean position of the Aleution Low, an area that has a high degree of variability, verification maps could be dominated by large differences in this region. To avoid this bias and determine where the significantly different errors occur the following statistical analysis was carried out.

For this application, a two tailed Student's t test was performed upon each grid point of the difference map. The t test is based upon the assumption of independent random samples from two populations which have the same variance.

The t statistic is computed by defining at every grid point within the analysis domain, n_f to be the number of values to compute the averaged forecast value and n_a defines the number of values used to obtain an average analysis value. Therefore n_a and n_f are simply the number of cases within the group (eg. all cases for the climatological group). The averaged forecast and analysis values at all grid points can simply be defined as;

$$\bar{X}_f = \sum_k x_{i,j,k} / n_f \text{ ,forecast of k cases;}$$

$$\bar{X}_a = \sum_k x_{i,j,k} / n_a \text{ ,analysis of k cases.}$$

The sample variances are computes for all grid points of each field as;

$$S_f^2 = \frac{1}{n_f - 1} \sum_k (x_{i,j,k} - \bar{X}_f) \text{ , forecast;}$$

$$S_a^2 = \frac{1}{n_a - 1} \sum_k (x_{i,j,k} - \bar{X}_a) \text{ , analysis.}$$

The actual t value at each point is defined as;

$$t = \frac{\bar{X}_a - \bar{X}_f}{S\sqrt{\frac{1}{n_a} + \frac{1}{n_f}}} \text{ , for all grid points.}$$

where

$$S^2 = \frac{(n_a - 1)S_a^2 + (n_f - 1)S_f^2}{n_a + n_f - 2}$$

Contours of the t values can therefore provide error patterns which illustrate those analysis minus forecast differences which have been objectively determined to be significant.

We have now defined a long term mean error pattern for our chosen synoptic scale. We define this mean error as being our systematic error which has all smaller scale, rapidly fluctuating errors "filtered" out by the averaging process. Of course the mean error pattern should be examined in terms of the physical reality of the pattern.

The objective minimum data period can then be determined by comparing shorter term error patterns with this mean pattern. Any shorter term or time scale error patterns which do not resemble or contain the mean error pattern characteristics may be defined as containing more random fluctuations than systematic signal.

The shorter term error patterns shall be defined by dividing the longer term, mean groups which defined our climatology pattern into smaller samples. The number of shorter term groups and the composition of each group is determined randomly. Random numbers shall be chosen to determine how many cases shall be included within each group. In order to insure a similar distribution of cases within each group another form of the t test shall be applied to evaluate the hypothesis that the mean of each group is not significantly different from the climatology mean. This will insure that no group is biased by the extreme or rare event.

The t test for this application is set up somewhat differently than the formulation described above. We define the mean of the climatological forecast field, say, as $\overline{X}$ computed as;

$$\overline{x} = \sum_i \sum_j x_{i,j} / N$$

where i,j are the grid points within the analysis domain, totalling N. The mean of the forecast fields for shorter term groups is computed in a similar manner and defined as $\overline{Y}$. Again N is the total number of grid points within the domain. The sample variance is computed for the shorter term group field as

$$S^2 = \frac{1}{n-1}\sum_i\sum_j (y_{i,j} - \overline{Y}).$$

The t value is then defined as;

$$t = \frac{\overline{Y} - \overline{X}}{S/\sqrt{N}}, \quad \text{for group k.}$$

This t value can be used to determine whether the shorter term group forecast and analysis fields are being effected by unusual case. Therefore the short term group is similar in amplitude to the climatological field.

Error patterns are computed, using the two tailed t evaluation for each shorter term group. As an example we may have, say, 20 cases of a 72 hour forecast and verifying analysis which may be averaged to form our "climatological" error pattern. Randomization may produce a set of short term groups distributed, say, as follows:

group 1	3 cases;
group 2	6 cases;
group 3	11 cases;

The hypothesis tests have also concluded that the amplitude of the mean of each group is similar. If the hypothesis is rejected the case may be re-randomized until all groups are similar.

The remaining task is to compare error patterns between all short term groups and the climatological group. This is accomplished by subjectively comparing the error patterns and by applying a variation of the pattern recognition techniques outlined by Somerville(1977).

Before detailing the comparison analysis it is better to define the remaining procedural concepts. As stated in Section 1 there are three possible cases which may exist at this stage of the analysis. These are illustrated using the example of the 3 groups above.

Case I

Case I may be defined when the resultant error patterns of the mean group and shorter term groups are defined to be similar. This would occur, say, if group 1, above which is constructed of 3 cases contains the same error pattern and centers existing in the mean group. Also group 2 (6 cases) and group 3 (11 cases) are also similar to the mean group.

To verify that the error pattern of group 1 (3 cases) was not obtained purely by chance, a number of groups containing 3 cases will randomly be constructed and the error patterns computed. This procedure is carried out on the shortest time scale which shows systematic error patterns. For example if 10 random 3 case groups are constructed we need, say 9, or 90% of the groups to have similar patterns to the original 3 case group. If this occurs we would accept the original 3 case group and conclude that our shortest term average contains systematic error components. This suggests for the chosen synoptic event the systematic error signal is very strong and no smoothing or averaging of random components may be needed.

Case II

Case II is defined when error characteristics are uniform for all but, say, the shorter time scales. For example, it may occur that the 11 case group and climatological group error patterns are similar but the 6 case group and 3 case group are not similar to the mean pattern. This would indicate that, perhaps the average of 6 cases is not sufficient to filter out smaller scale random errors. This would be verified by forming a number of 6 case groups to verify the initial 6 case error pattern.

Because the 6 case group does not match the mean error pattern and the 11 case group does, it may be possible that the optimum data period be between these two groups. Therefore, randomized 7 case, 8 case, 9 case and 10 case groups are constructed. Error patterns for these groups are computed to be compared with the 11 case and mean group. This will determine the final data period necessary for filtering out random errors. It may happen that the 9 case group matches the mean error but the 8 case does not. Therefore the 8 case error pattern would be subjected to the randomization verification test. If a sufficiently large number of random 8 case groups are similar to the original and not matching the 9 case, 11 case or mean groups we conclude that at least 9 cases must be averaged to sufficiently show systematic errors. If the randomization test does not satisfactorily verify the 8 case pattern the process would be repeated.

Case III

Case III exists when no shorter term groups match the long term mean group. In the event of this situation a number of items may be concluded. It may be that the chosen synoptic event or data fields are not well defined in the beginning, containing many scales of motion. In this case, this particular analysis may not be able to sufficiently smooth any one scale. If the synoptic event is well defined and the data has been checked we may conclude that for this particular situation it is impossible to filter out random error components. This would imply that any verification studies using the model and synoptic event may require more use of unconventional methods

We may now return to the methods of comparing the resultant error patterns for the various groups. The easiest way and most straight forward approach is to simply subjectively compare the patterns. This is very adequate when patterns are significantly different as often occurs.

Another comparison approach which is utilized here is a limited application of the Pattern Recognition techniques outlined by Somerville(1977). The technique is applied as a tool for reduction of the dimension of the resultant error patterns. The error patterns which are computed t values may be defined as

$$T = f(x,y,z,tau,g)$$

where x,y are the horizontal space dimensions

z is the vertical level

tau is the time or forecast interval

g is the group number

We can set the following

$$tau = n\Delta tau$$

$$x = j\Delta x$$

$$y = i\Delta y$$

Therefore for a given time forecast interval, and height (pressure level) we have

$$T = f(i,j,g)$$

We can therefore produce a pattern matrix defined as the maximum and minimum centers (defined as i, j coordinates) for a specific group, (g). The various pattern matricies may be superimposed into one final matrix containing i,j coordinates for maximum and minimum error centers for various groups of a given forecast and height. This allows definition of features or error centers which are common to groups or not common to groups. The final step is to decide upon a tolerance where a feature or error center may be, geographically lined up between the mean group and the shorter groups. This tolerance was chosen to be the contours of the significant t values of the mean pattern. This allows comparisons between groups and an indication as to how error centers compare for various time scales.

2.3 Analysis Summary

The analysis is summarized as follows:

i) A specific synoptic event is chosen for analysis;

ii) A model and verification analysis climatology is calculated;

iii) Individual cases are randomly grouped such that the number of groups and the number of cases per group varies;

iv) The distribution of case between groups is objectively analyzed to insure no group is biased by a case or extreme event;

v) Error patterns are calculated for all groups using standard differences and a two tailed t analysis;

vi) Spatial error patterns are subjectively compared and analyzed using a limited application of pattern recognition concepts.

3.0 Application

This procedure was applied, in full, to a verification of the Shanghai Low synoptic event. 30 cases were chosen over the winters of 1977-1978 and 1979-1980. 24 and 72 hour forecasts were analyzed using the methodology described in Section 2.

24 Hour Forecast

The 24 hour forecast gathered for the analysis were divided into 2 time periods. Period 1 consists of 24 hour forecasts made on the time the storm was initially chosen as an event. The second period is 12 hours later than period 1. It is recommended to stratify the data in this manner, giving possible indications of how a model might handle the initial cyclogenesis of the event.

The initial period analysis is described here. The purpose of the application was to illustrate and test the analysis technique, not to determine or document any errors within the model.

The original 30 cases were randomly distributed into groups two times. The following groups resulted:

1) 3 cases;
2) 4 cases;
3) 7 cases;
4) 8 cases;
5) 11 cases

Figure 1 illustrates the averaged forecast and verifying analysis for each group including the average of all 30 cases or mean group. Figure 2 illustrates the error pattern, constructed of t values, for the mean group. Figure 3 shows the resultant error patterns for the remaining shorter term groups. The t maps were constructed from a difference of analysis minus forecast. Therefore negative values are associated with analysis pressures which are lower than the forecast

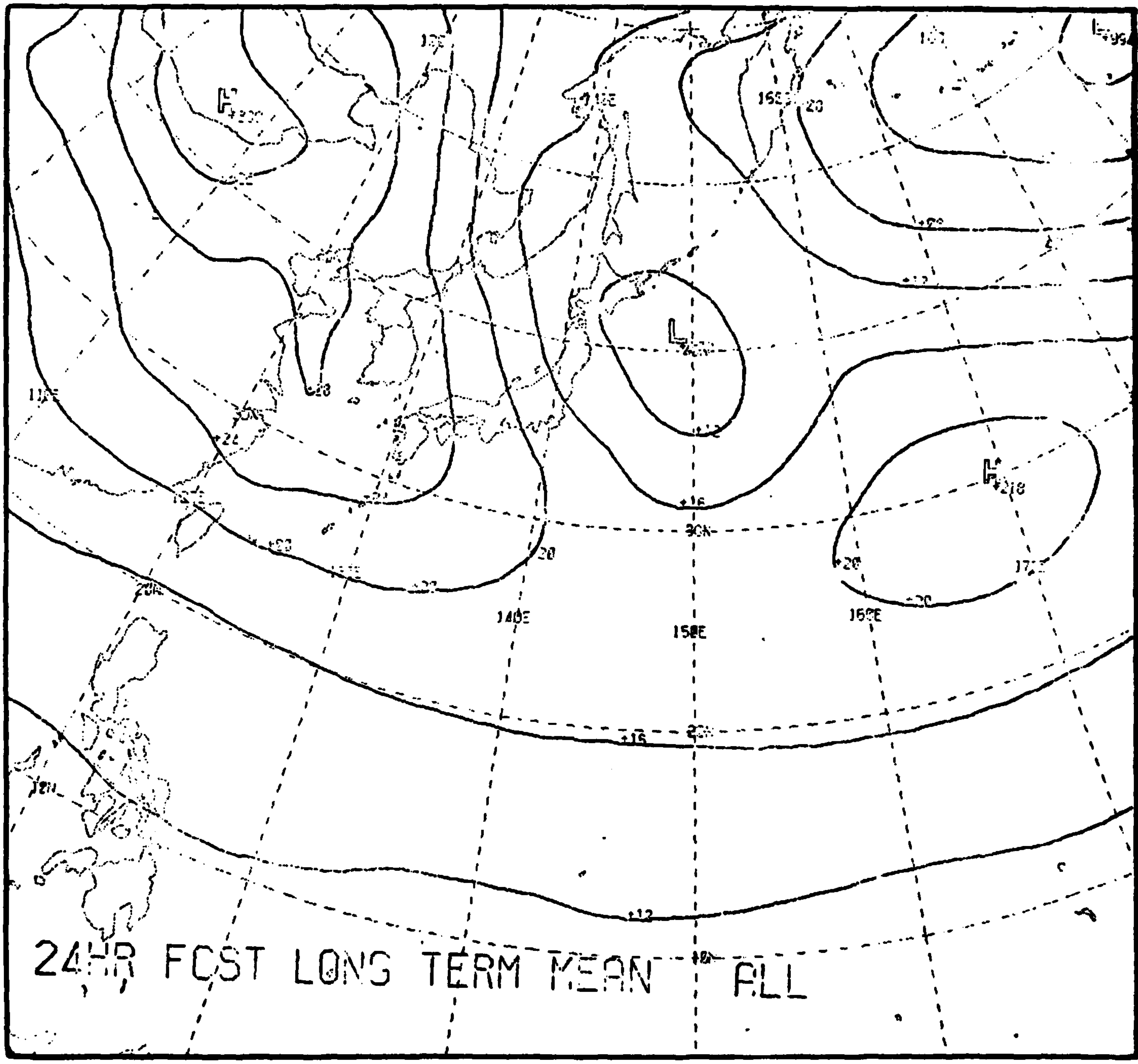
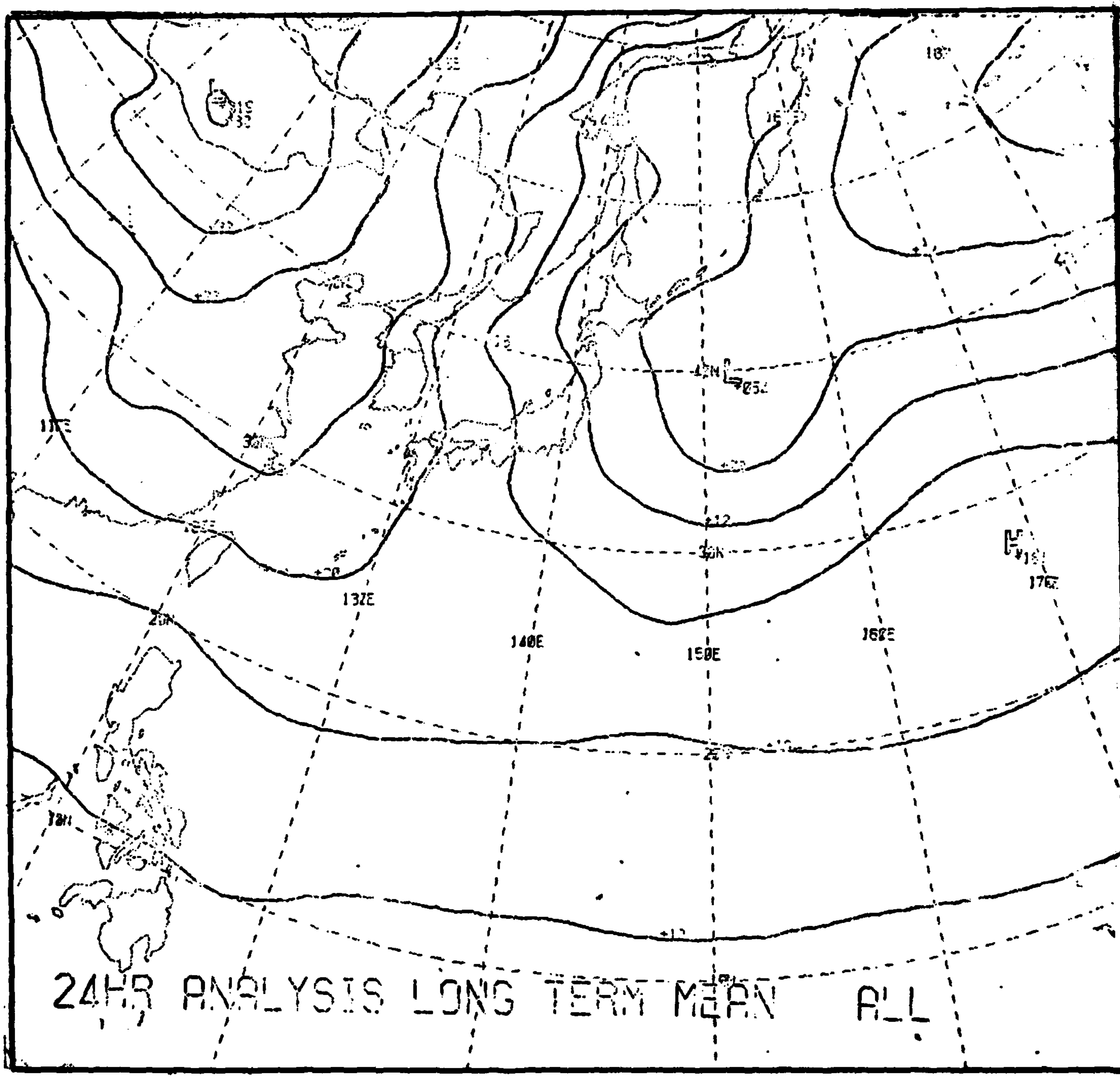


Figure 1 Averaged forecast and verifying analysis fields for the long term group and selected short term groups. Surface fields.



24HR ANALYSIS LONG TERM MEAN ALL
130E
140E
150E
160E
170E
30N
20N
10N
+12
+20

Figure 1b

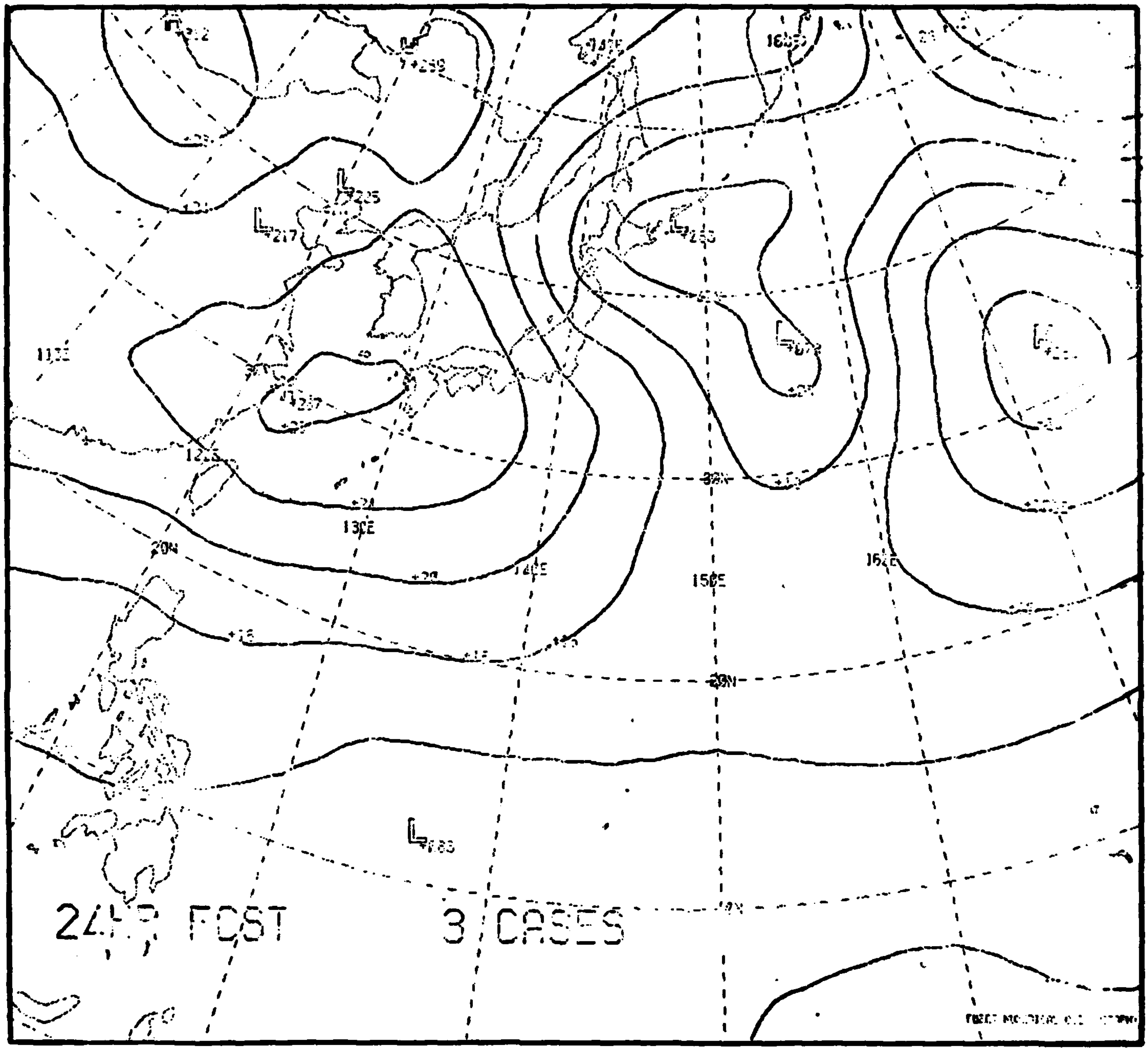


Figure 1c

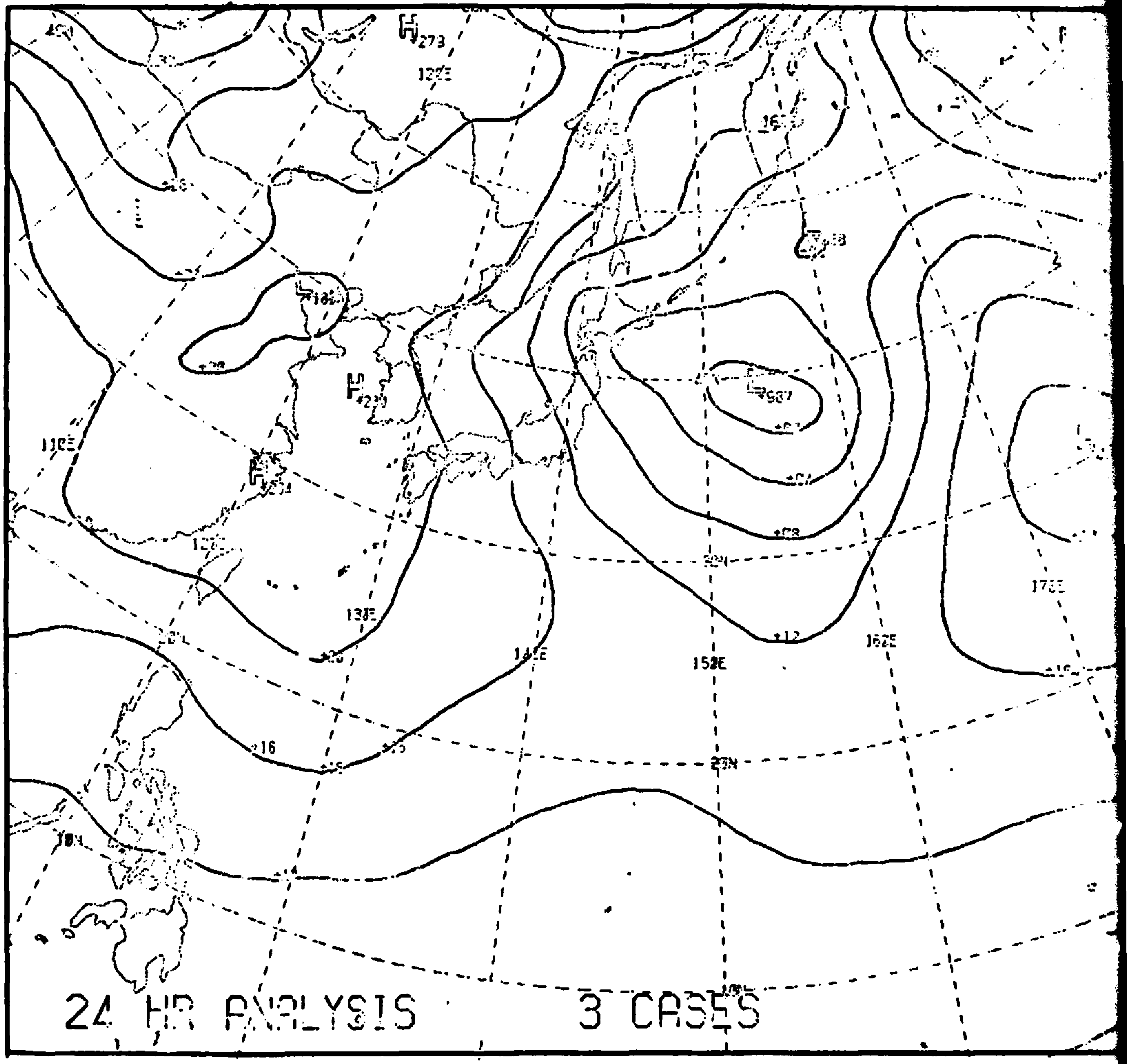


Figure 1d

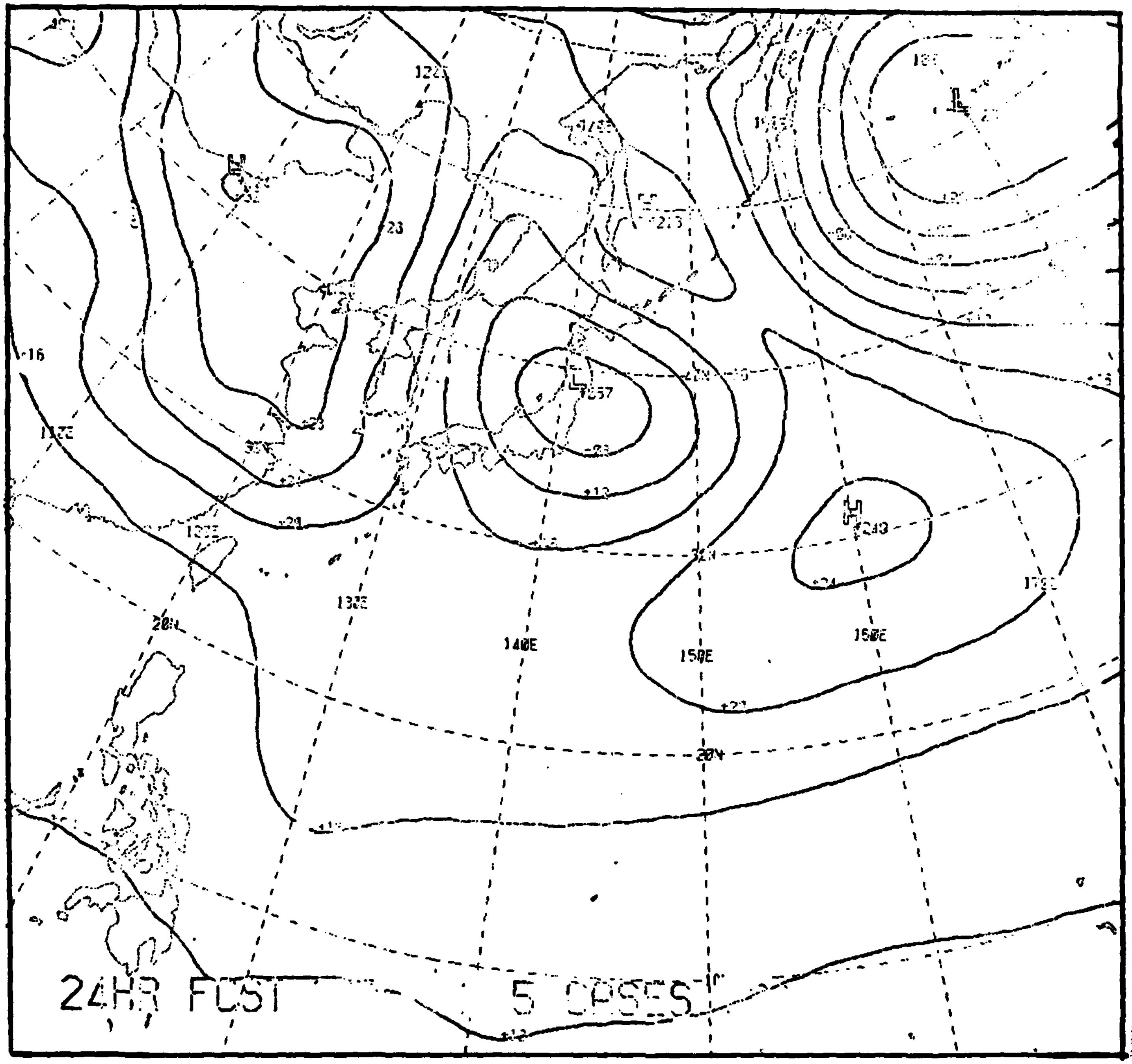


Figure 1e

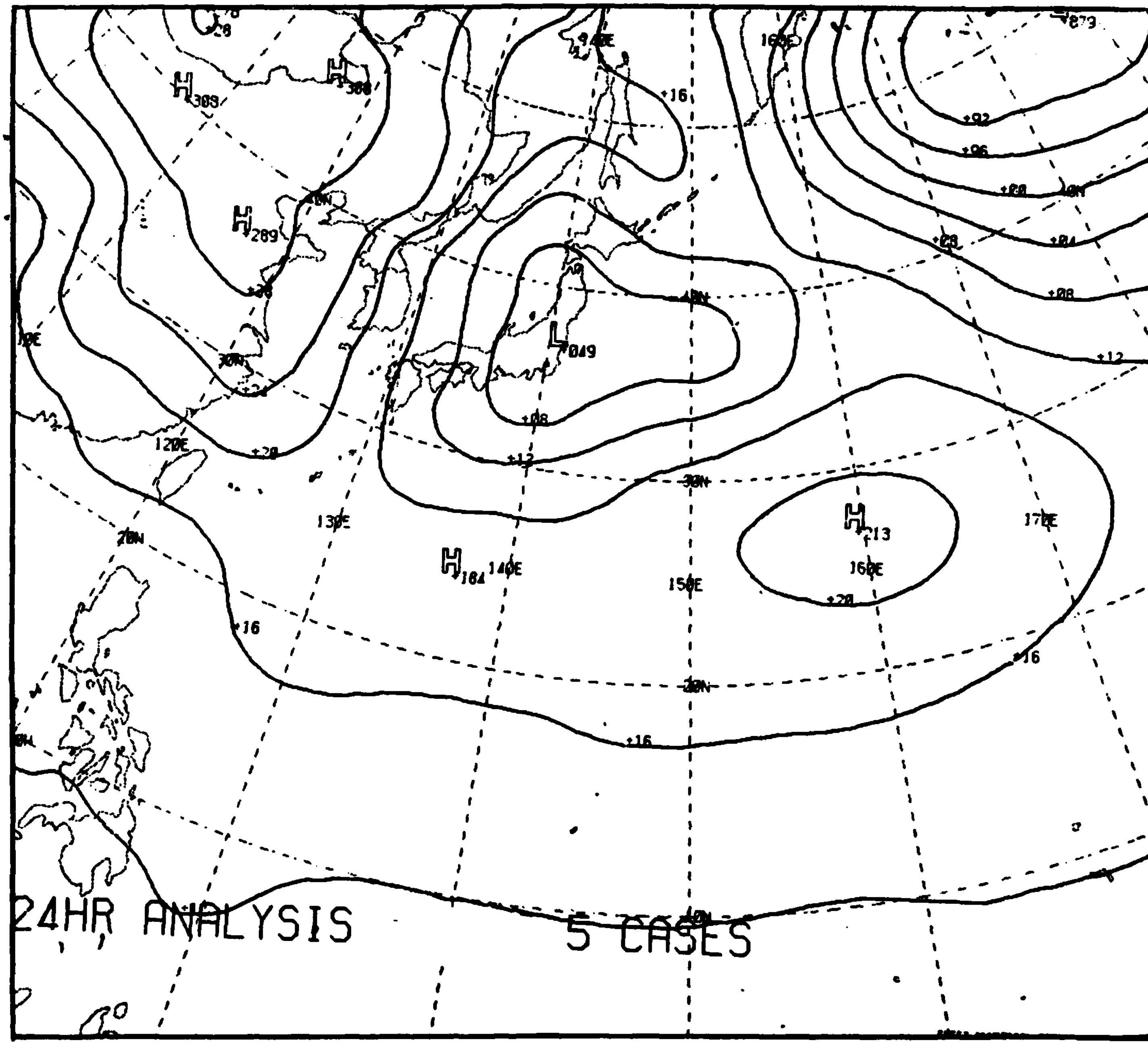


Figure 1 f

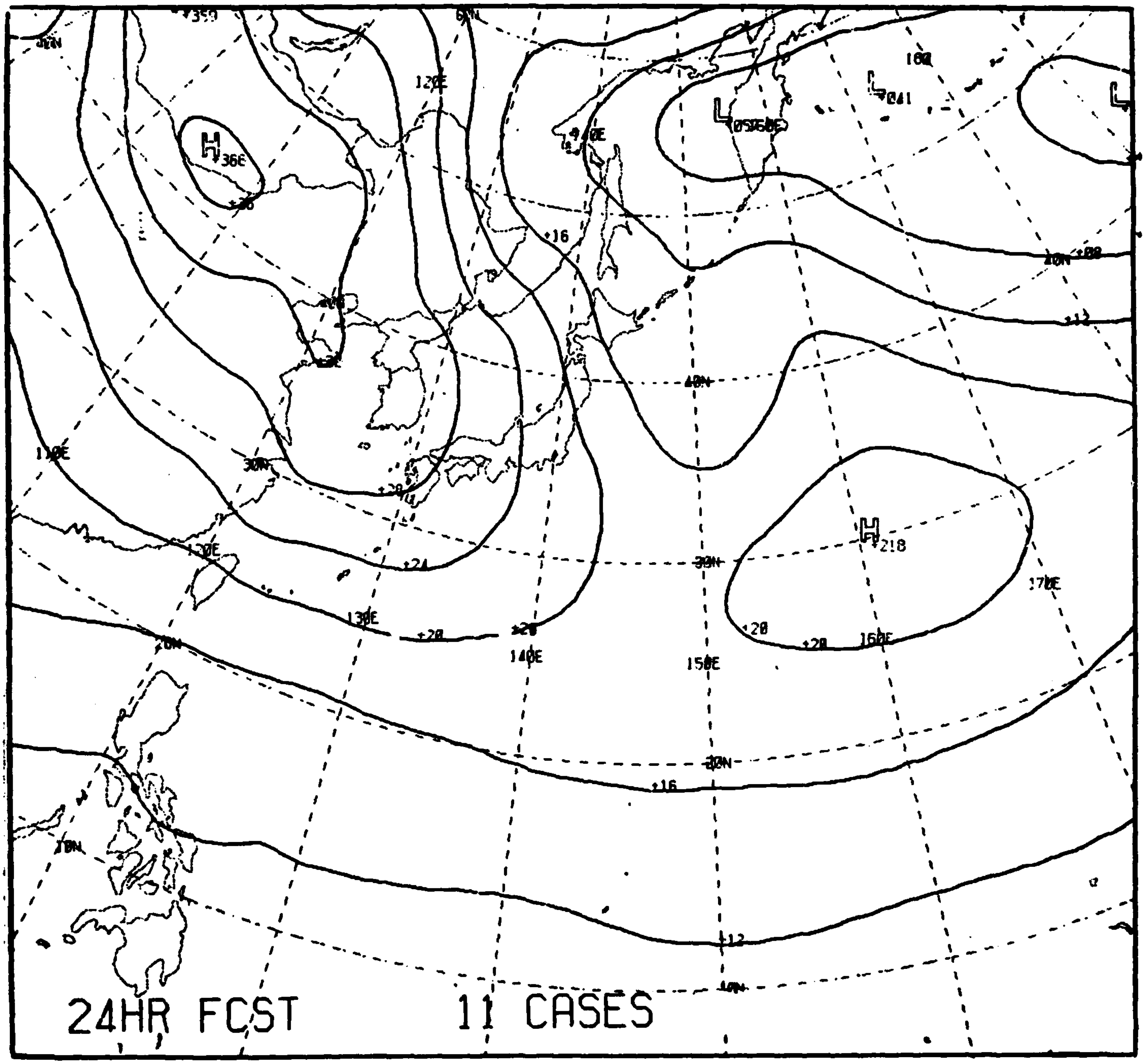


Figure 1g

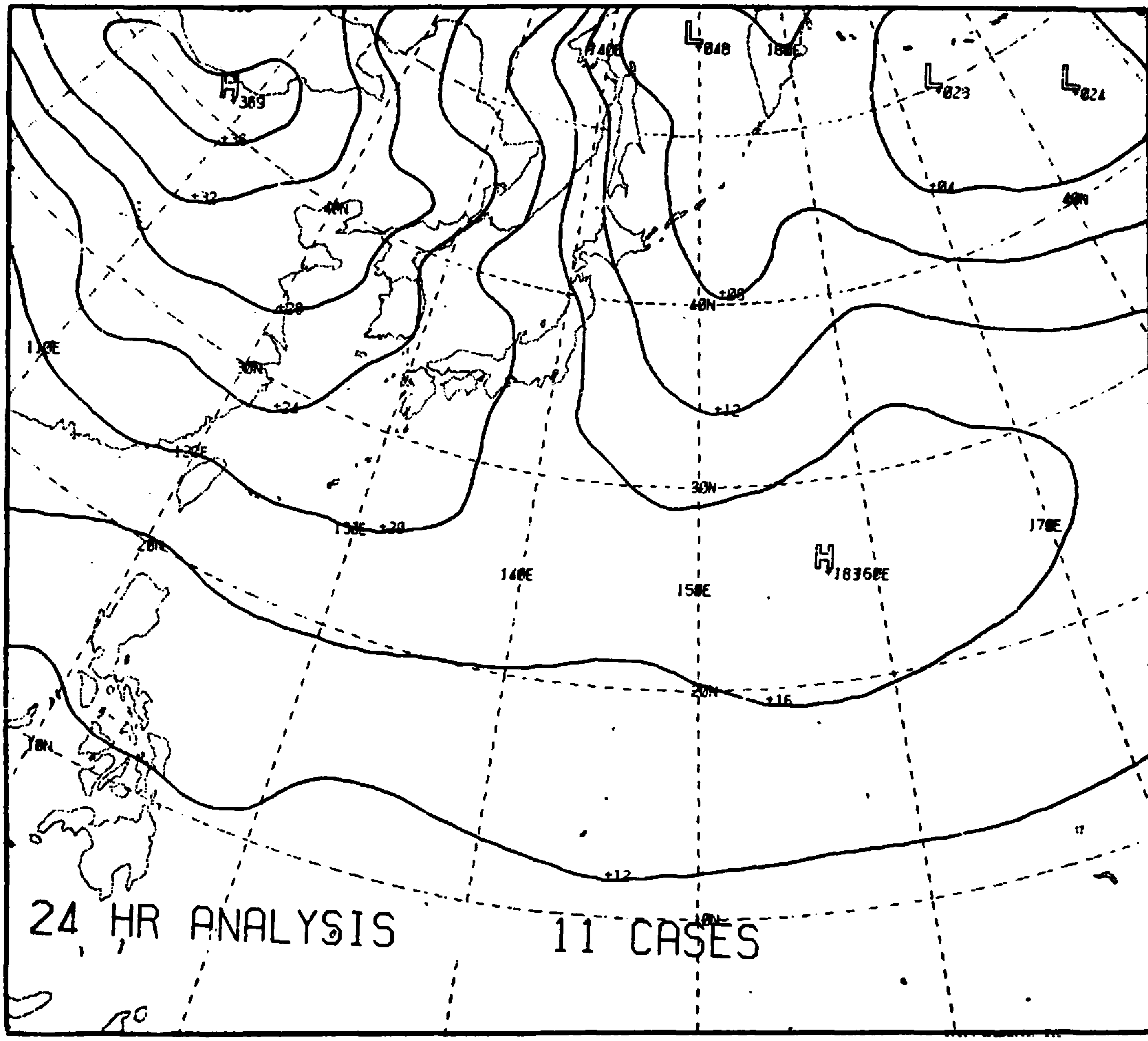


Figure 1h

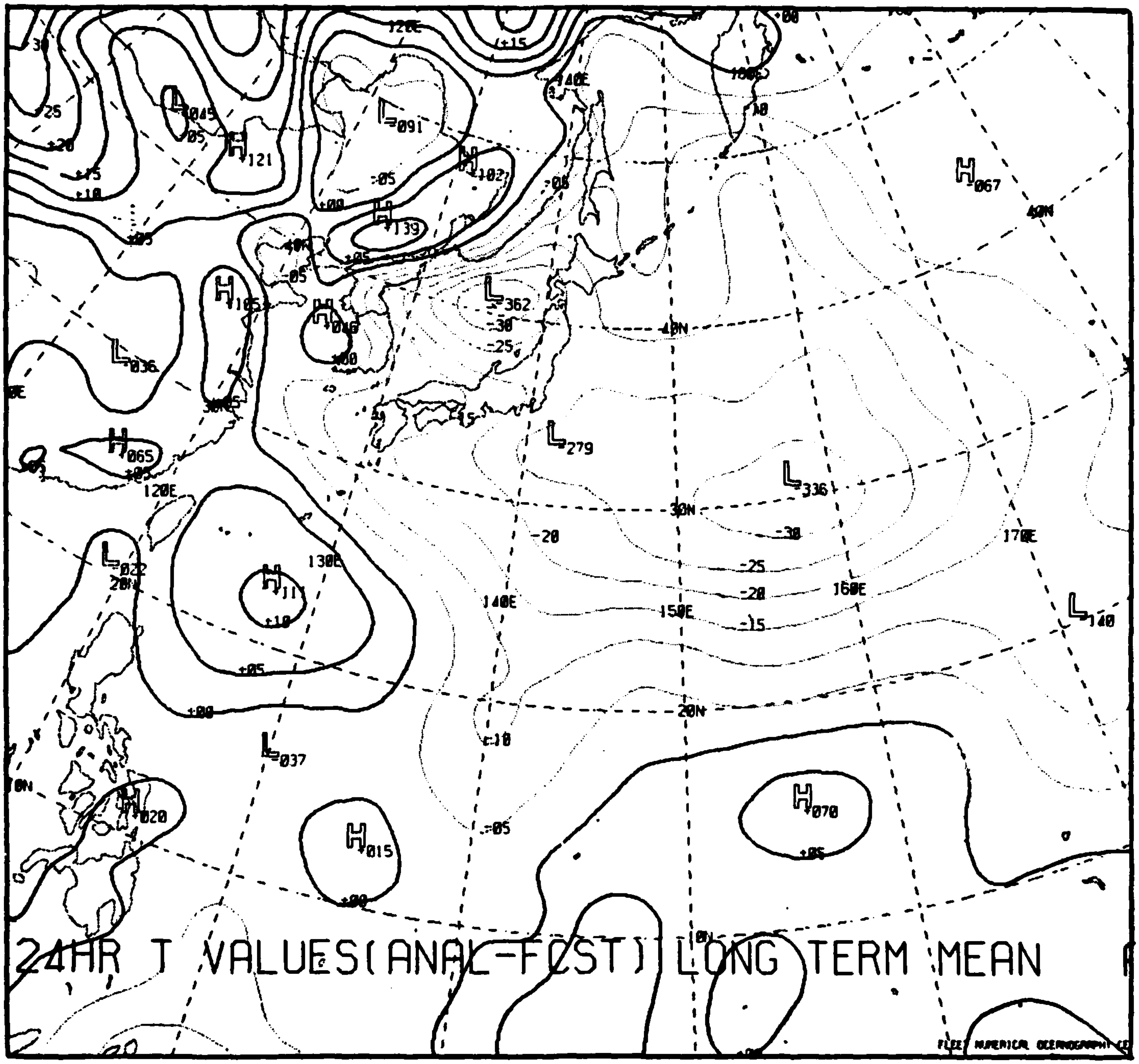


Figure 2 Error pattern constructed of t values for the long term mean group. 99% significance value = 2.3

values. The significance of the t values are presented in the figures.

Before these final groups were chosen, the t analysis used to insure that no group was biased was carried out.

Figure 2 shows a relatively smooth error pattern over the main region of interest. Noisy patterns over the Asian Continent have no relevance here. It should be noted that the great majority of Shanghai Lows form as a wave on the polar front which usually intersects the China coast between 25-30 N latitude. The mean error pattern is dominated by two centers of significant negative errors (analysis lower than forecast). The errors are concentrated over the Sea of Japan and the Pacific Ocean extending southeastward from Japan. The third dominant feature, though less significant, is a center of positive error in the region where the events climatologically form. A possible interpretation of this error pattern would be at 24 hours the model tends to predict lower pressures in the general region of initial cyclogenesis rather than move the storm along. The large centers of negative error occur in regions where a storm of this type would climatologically travel. However the model has tended to over predict the deepening of the storm in the area of initial formation and under predict the propagation and movement of the storm for the 24 hour period. Notice that the error pattern of figure 2 contains only the 3 described significant error centers providing a smoothed error pattern within the region of interest. The noisy patterns discussed around the region are not nearly significant as the three described areas.

We may now turn to the randomized groups shown in figure 3. Figure 3a contains the 3 case error pattern. This pattern is very disorganized. There are numerous error centers which exhibit some statistical significance. The three main centers described in relation to Figure 1 are barely discernable. This pattern clearly represents a case where the smaller scale random fluctuations are very much present in the error field. It would be very difficult to obtain a clear understanding of an error pattern such as this, which was simply obtained in Figure 2.

Figure 3b(4 cases) exhibits a smoother pattern than figure 3a. However there are still a rather large number of statistically significant error centers dispersed throughout the region of interest. The distribution of positive and negative errors is more uniform than observed in figure 3a, however.

Figure 3c(7 cases) is beginning to resemble the mean case. The main error centers are very visible and significant centers. The main difference is that the larger number of significant centers are clustered within the areas described in relation to figure 2.

An inspection of figures 3c and 3d shows that at 8 and 11 cases, the error patterns closely resemble the long term mean error centers. The three main centers are easily discernable however the significance of the positive error near the South China coast has been further reduced.

Figure 4 illustrates the application of the pattern recognition technique showing maximum and minimum centers of the t analyses (figures 2,3b,3c and 3d) plus the 99% significance contour for the long term mean(figure 2). The purpose of the pattern recognition application is to measure how well error centers coincide between the long term and shorter term groups. The figure is constructed by first taking the t analysis of figure 2 and defining the significant error centers only. The domain of the analysis was reduced to avoid the noisy area over the asian continent. Therefore, looking at figure 2, we see two centers of significant differences. Both of these centers are negative and extend southeastward from Japan. These centers are marked by the i,j grid point, (x), and the data value to the right of the point on figure 4.

An inspection of figure 3b shows numerous error centers within the analysis domain. These centers are also marked on figure 4 by the i,j grid point, (x), and the data value is placed above the point. We see in figure 4 that one error center of figure 3b coincides exactly with the error center of figure 2 while another center of figure 3b is contained within the 99% significance contour of figure 2. However there are numerous other error centers of figure 3b which are dispersed randomly throughout the analysis domain.

The same procedure is followed, using figures 3c and 3d respectively. The error centers, defined by low and high symbols on the t analyses of figures 3c and 3d are marked on figure 4. Figure 3c, containing 7 cases, error centers are plotted to the left of the i,j grid point on figure 4 while figure 3d, containing 8 cases, error centers are plotted beneath the i,j grid point.

An inspection of figure 4 shows that one 7 case error center coincides with the long term or climatological error center(defined in figure 2) of 3.36. However, similarily to the 4 case group(figure 3b) there are numerous other error centers which seem to be randomly dispersed throughout the region of interest. Two 8 case(figure3d) centers coincide with the long term centers(figure 2) with only one 8 case error center lying away from the centers defind by the climatological error pattern of figure 2.

This analysis points out groups which have similar error patterns or error centers with the long term or climatological error pattern. Error centers of the short term groups are said to coincide with the long term groups when they lie within the 99% significance contour defined by the t statistic of the climatological error pattern.

Based on the above analysis we would conclude that at 8 cases the random fluctuations are sufficiently smoothed. This allows a clear identification of the true error pattern associated with the chosen synoptic event. Therefore we would define 7 cases as the cut off value for

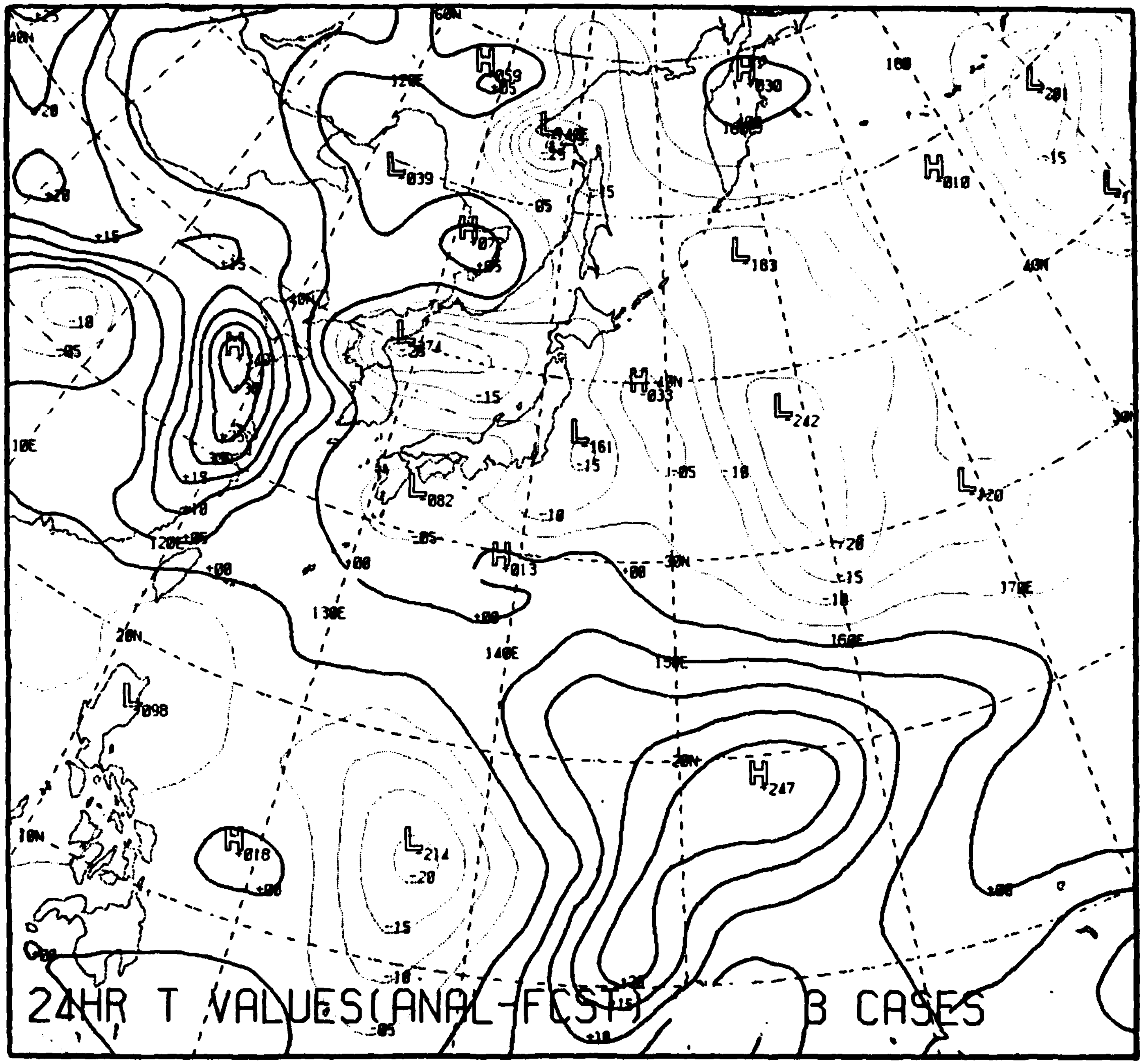


Figure 3 Error patterns for short term groups

Figure 3a 99% significance level = 3.7

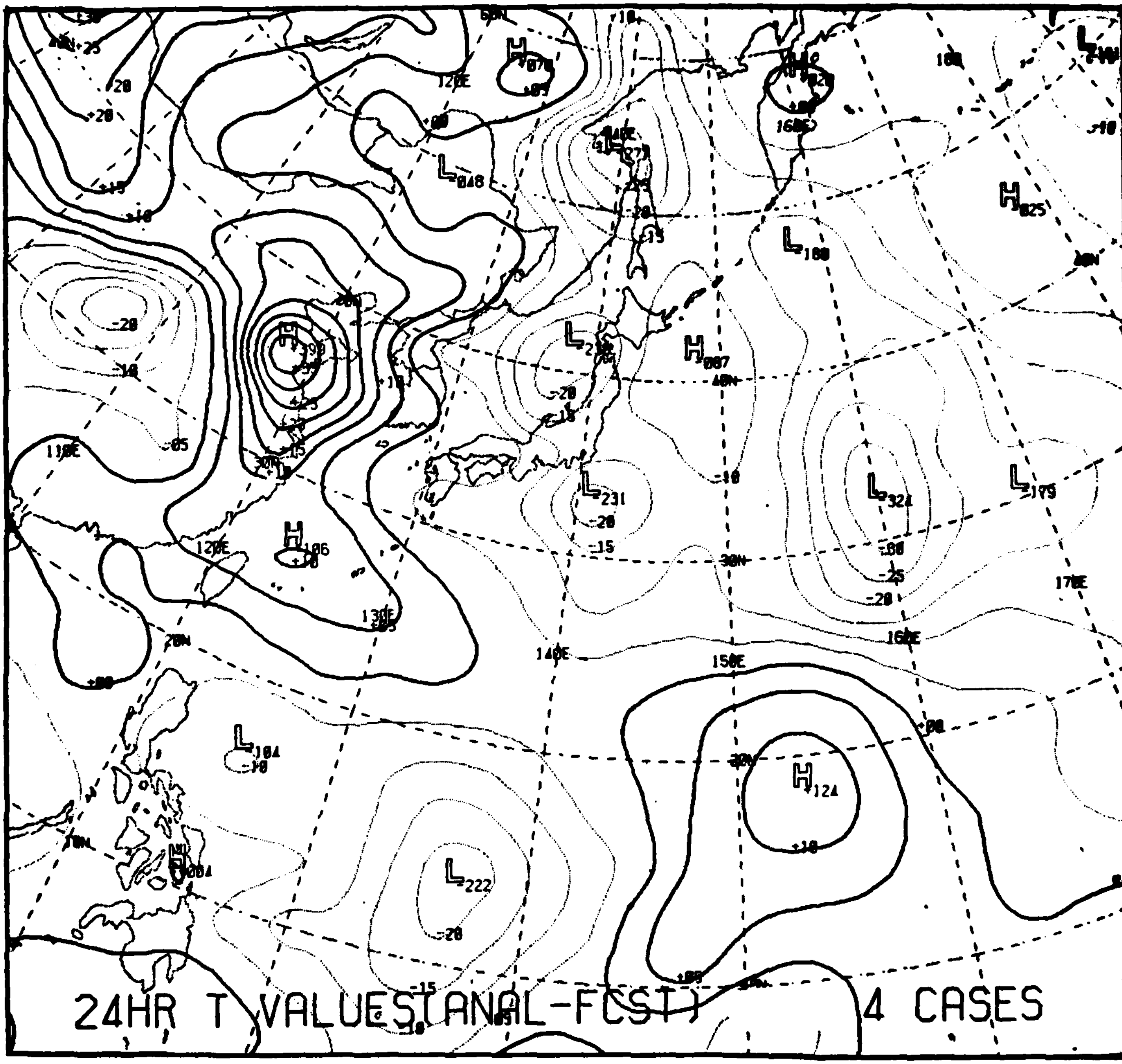


Figure 3b 99% significance value = 3.1

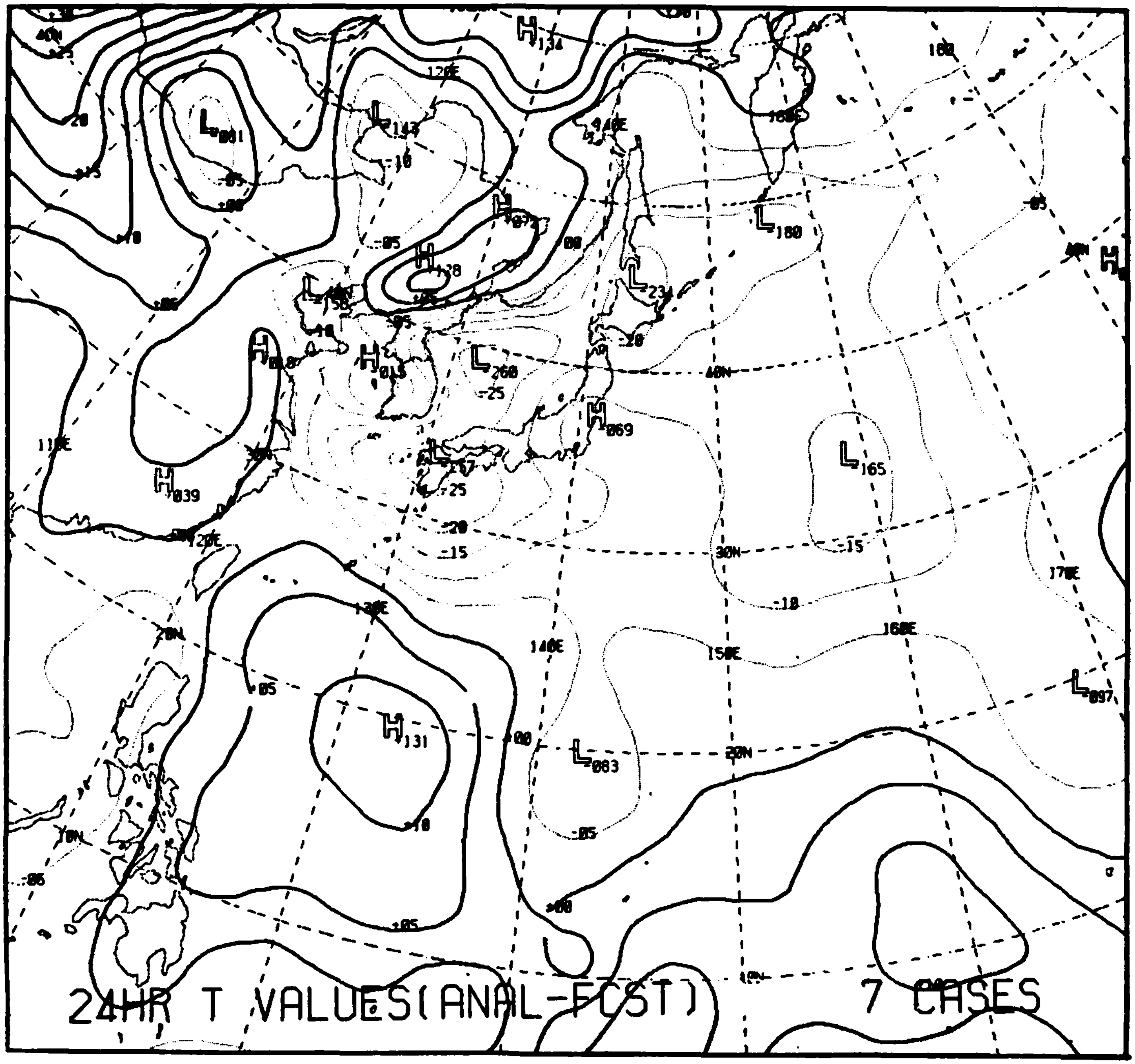


Figure 3c   99% significance value = 2.7

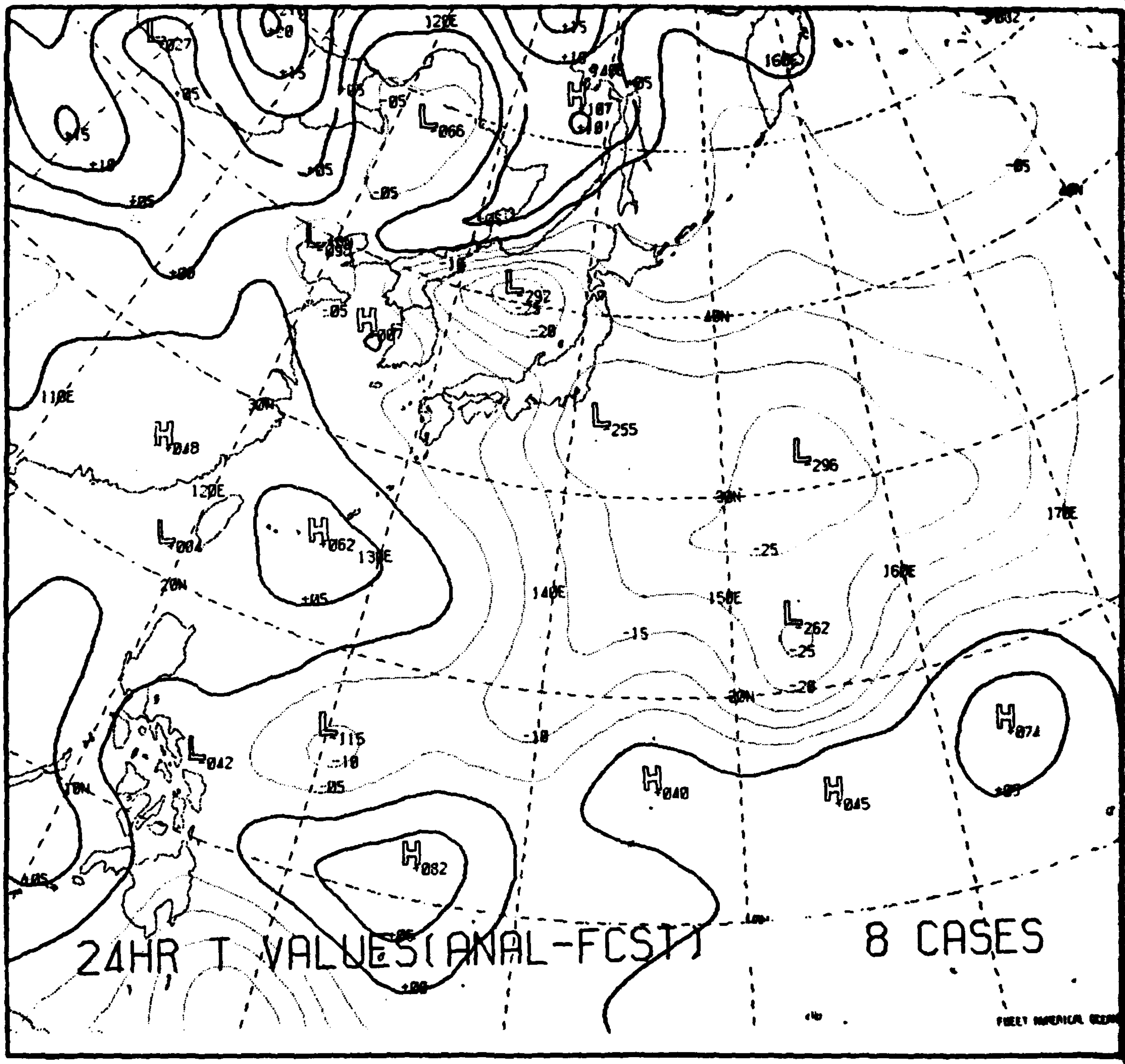


Figure 3d   99% significance level = 2.6

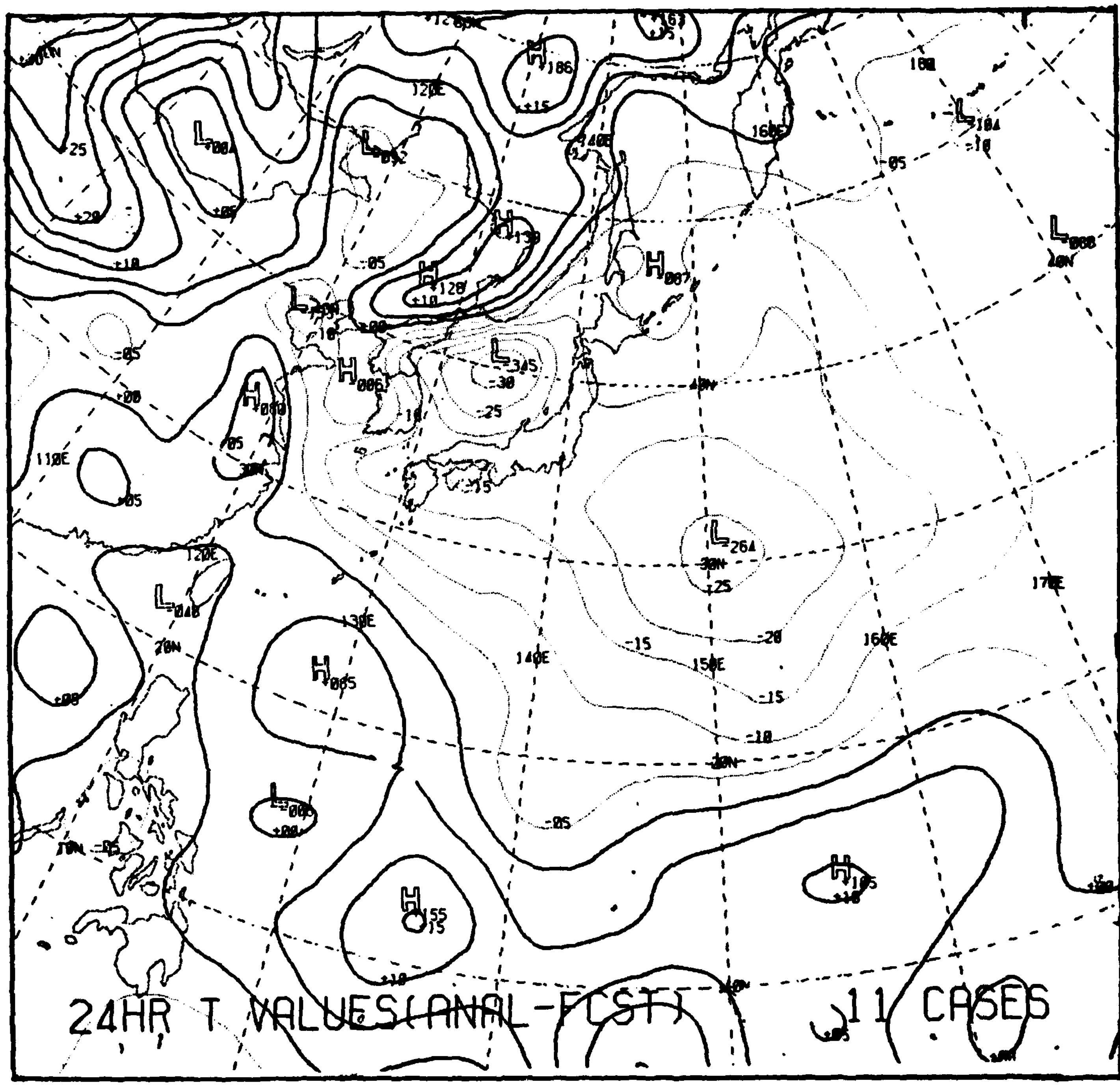


Figure 3e 99% Significance level = 2.5

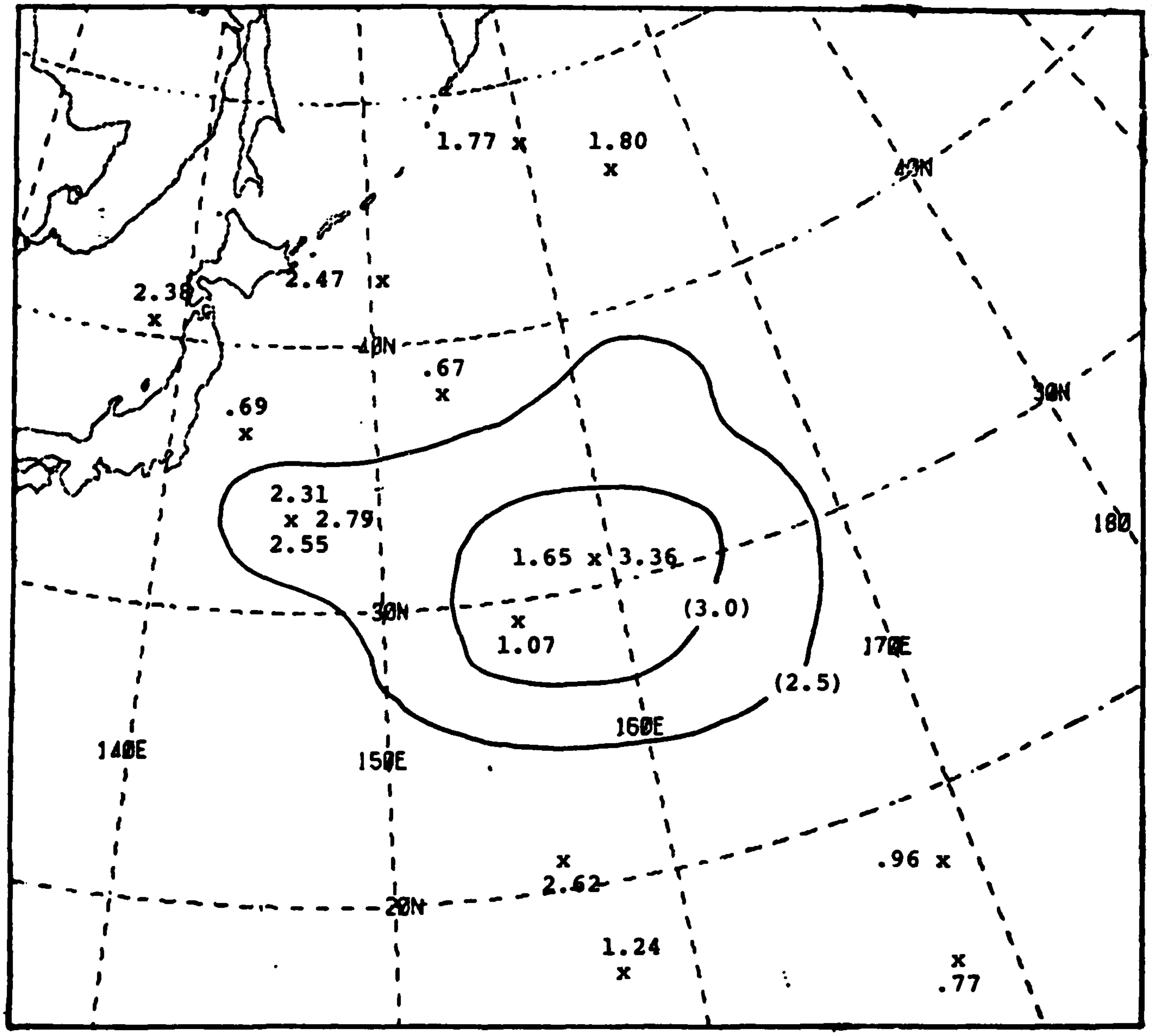


Figure 4. Maximum and minimum centers of the t analysis. All numbers are negative. Contours represent the 99% significance level of figure 2. Figure 2 centers are right of the point, figure 3b above, figure 3c to the left and figure 3d below.

which we would want to verify the pattern by randomly computing a number of 7 case groups. When we perform this function we are trying to identify significant error centers which are not corresponding to the long term group in the Sea of Japan and the Pacific Ocean extending southeastward from Japan. Ten randomly chosen 7 case groups were constructed All cases were part of the original 30 cases. Inspection of the random groups shows that all groups contain a sporadic number of error centers in the negative region of the North Pacific Ocean. This validates the original 7 case pattern and the 8 case group as the minimum data period for averaging cases together to sufficiently smooth out random fluctuations.

72 Hour Forecast

The 72 hour forecast cases were stratified by time in the same manner as the 24 hour forecasts. The following smaller term random groups were constructed.

1) 3 cases;
2) 4 cases;
3) 5 cases;
4) 7 cases;
5) 8 cases;
6) 11 cases

The respective forecast and analysis fields are shown in figure 5. Figure 6 shows the error pattern (t values) for the long term mean group of the initial time period cases. There are 30 cases in all.

The error pattern of figure 6 is dominated by 2 large cells. Positive errors (forecast values greater than analysis values) are dominant in the East China Sea similar to the 24 hour case of figure 2. The error center of this region is much more significant however. The second large error center covers the Pacific Ocean, east of Japan and extends northward towards the Kamchatka Peninsula and westward into the Sea of Japan. Again this pattern is quite smooth and

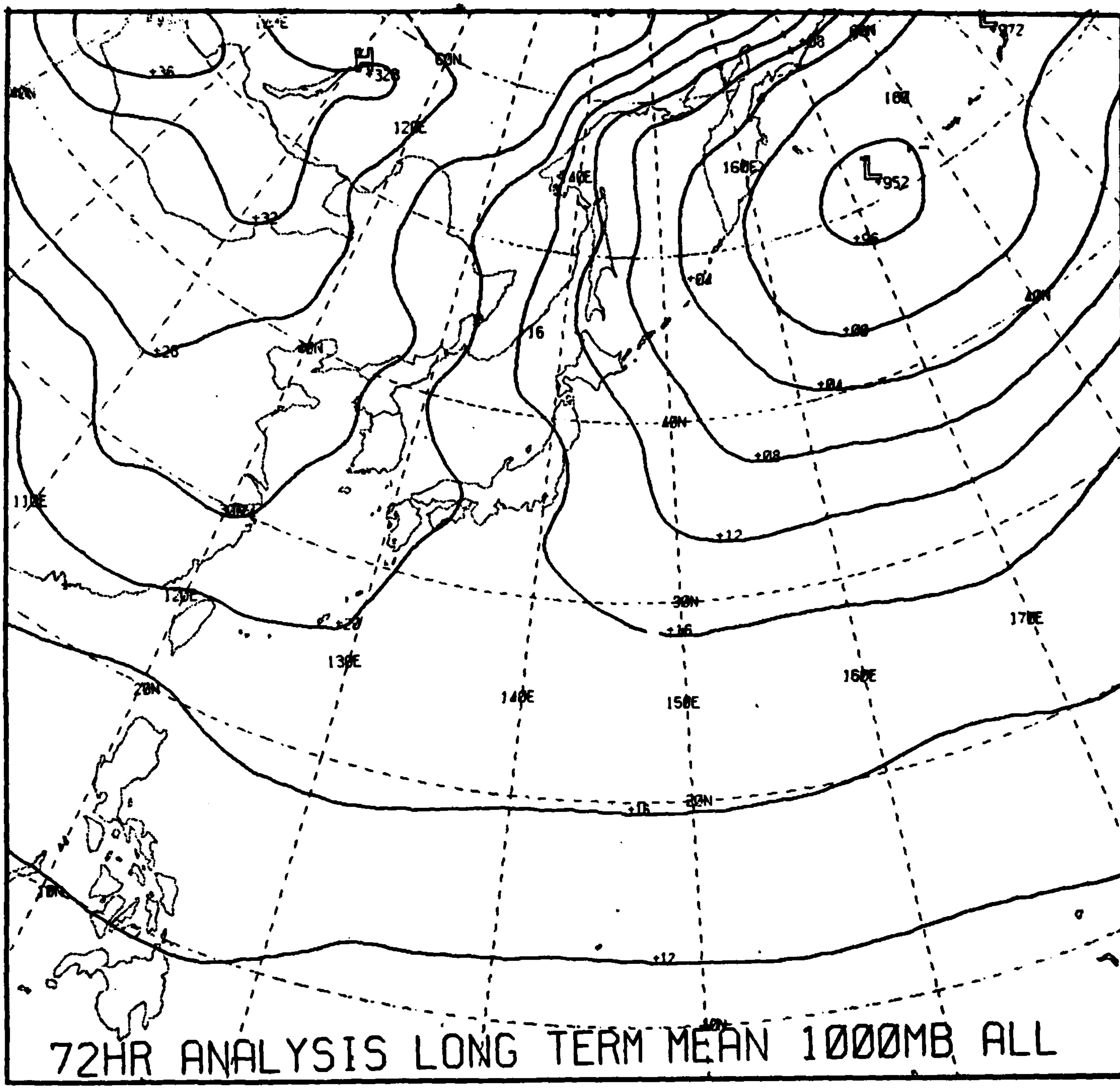


Figure 5 Averaged forecast and veryfying analyis fields for the long term mean group and selected shorter term groups.

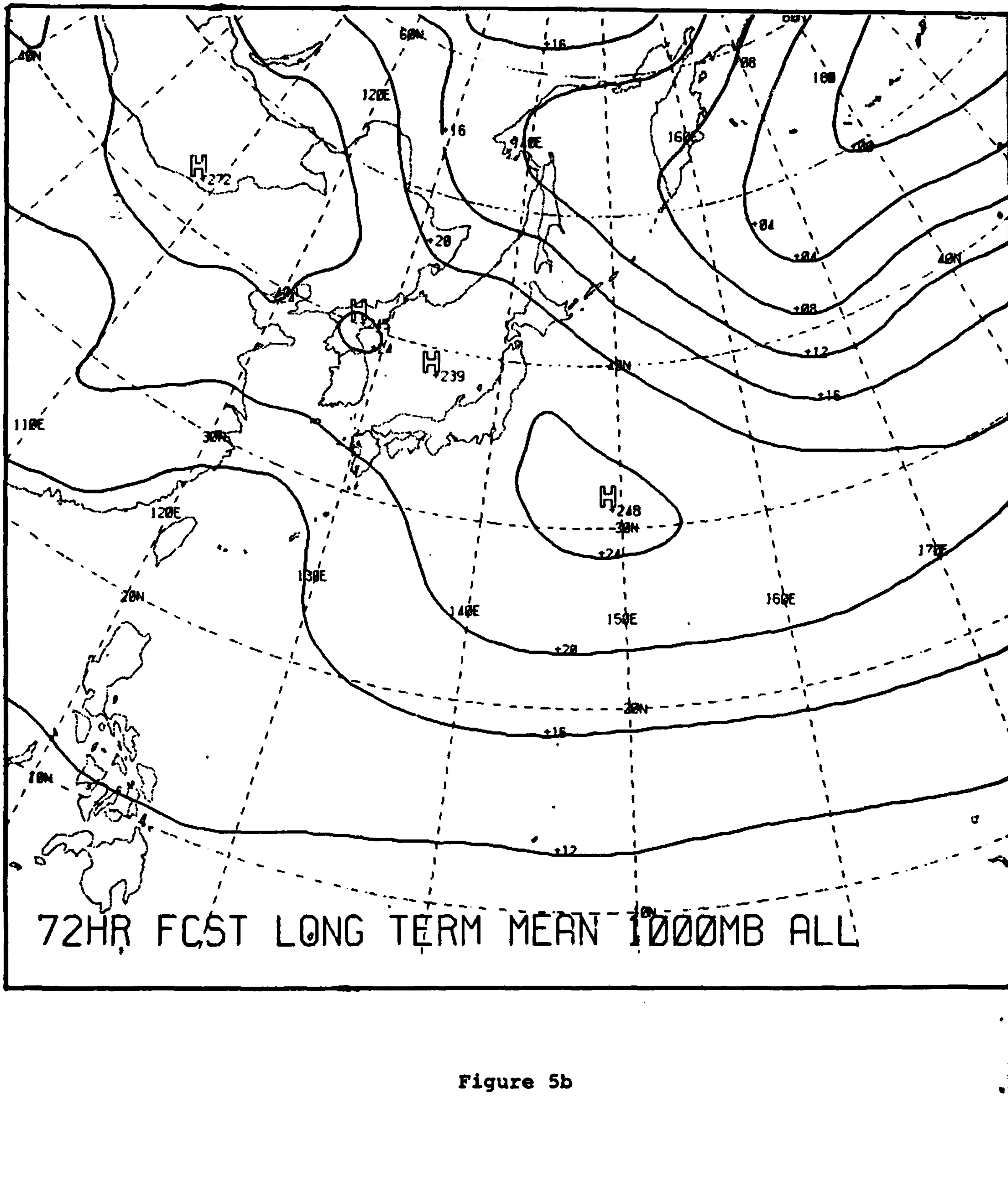


Figure 5b

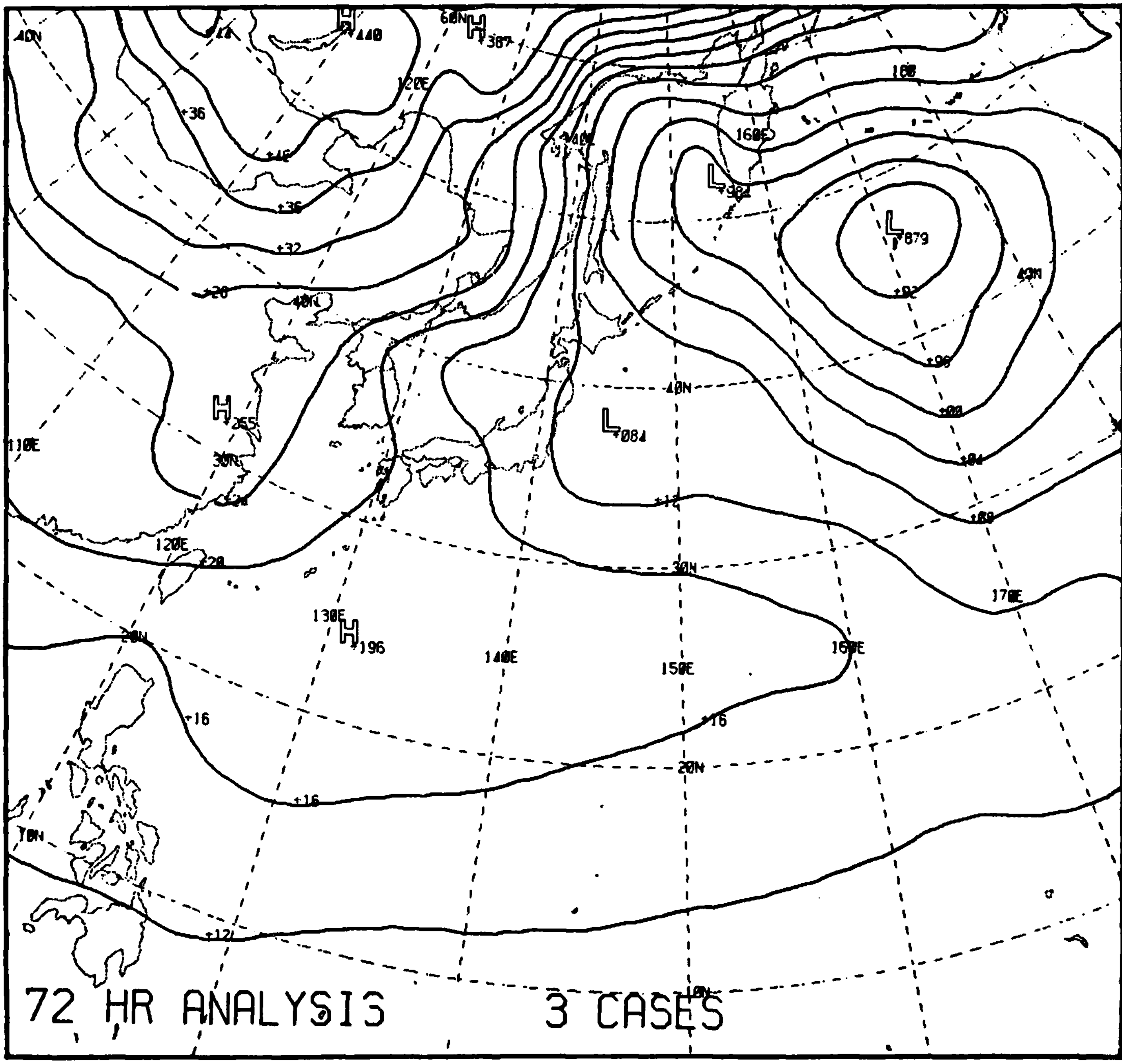


Figure 5c

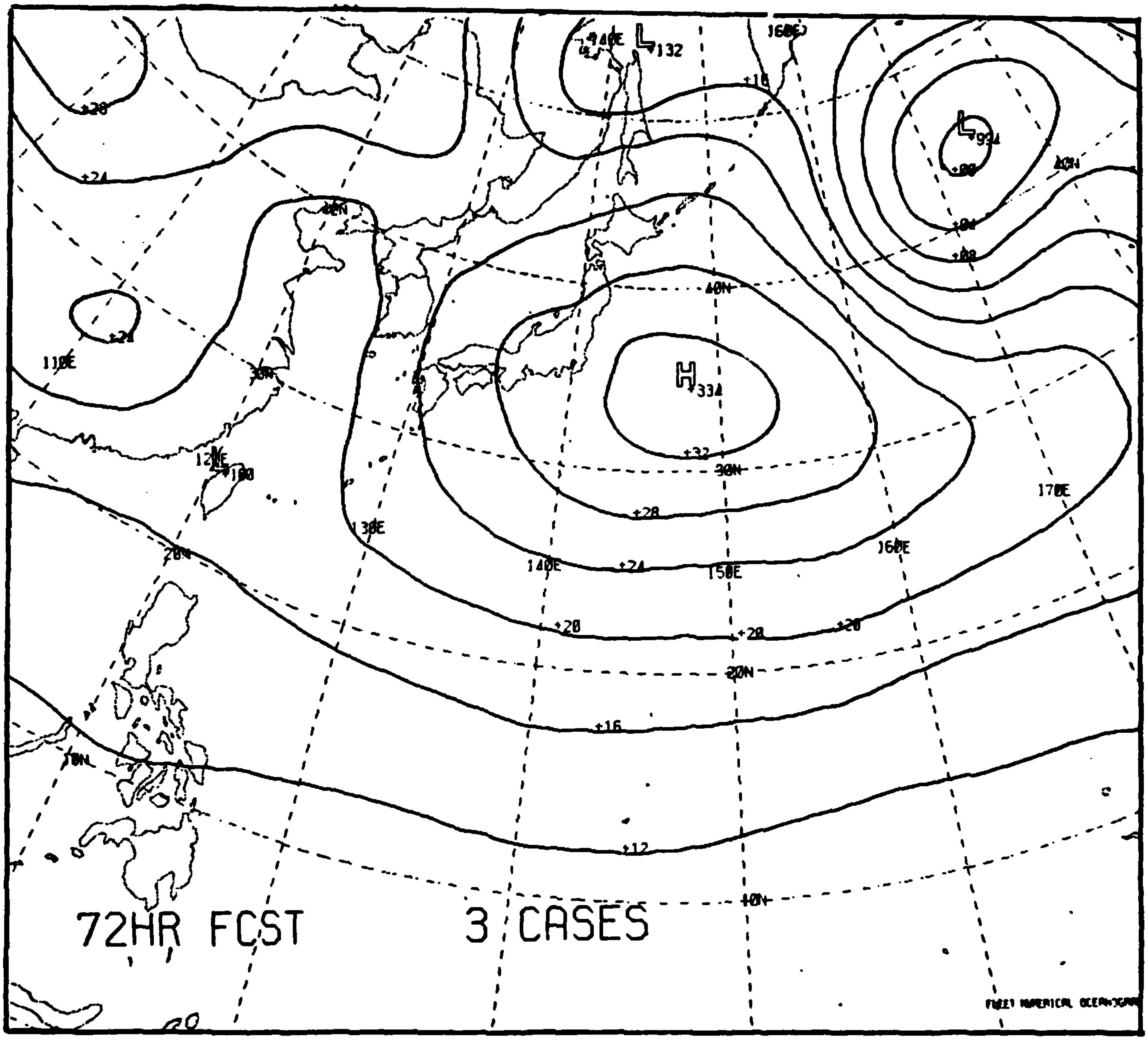


Figure 5d

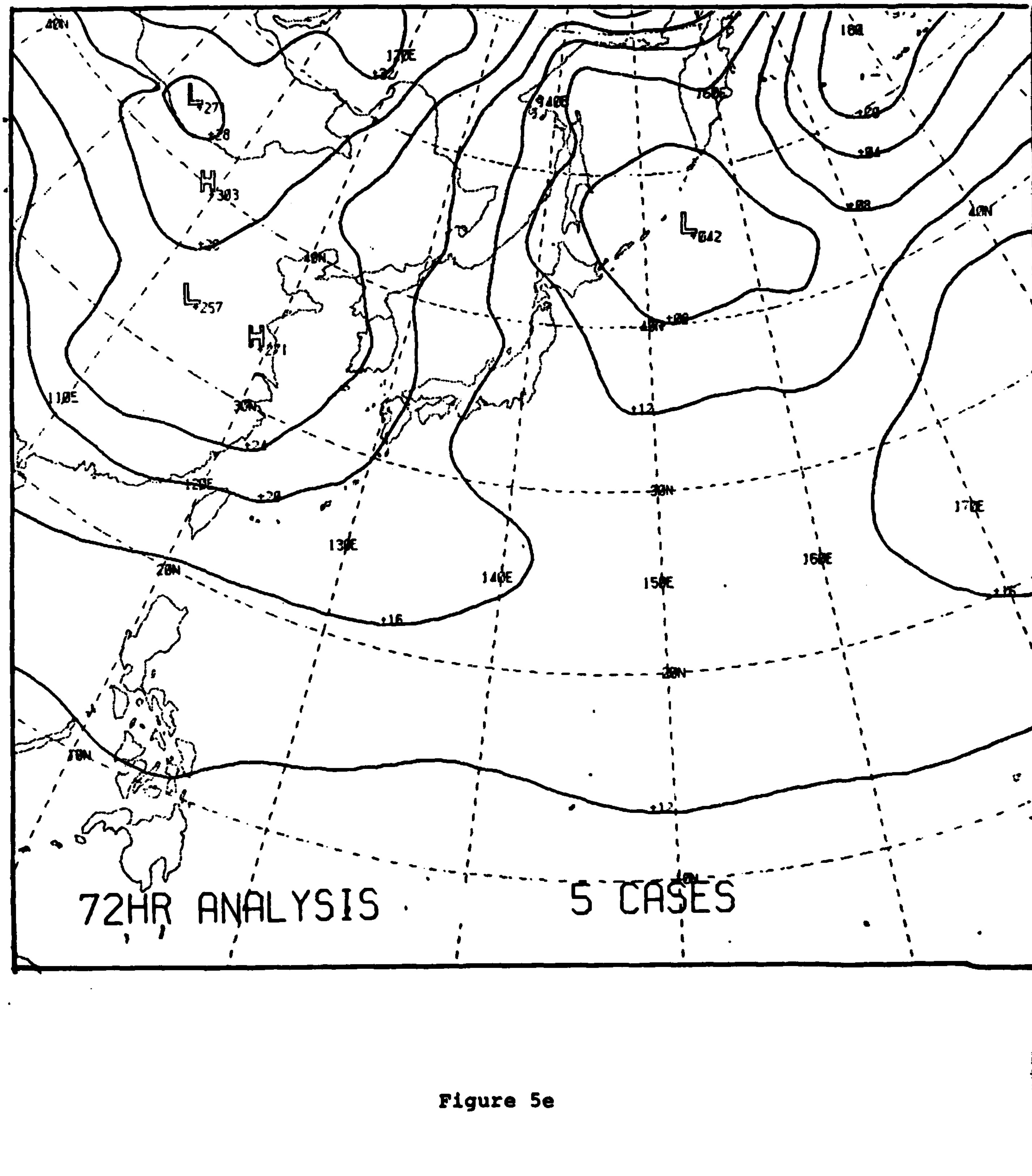


Figure 5e

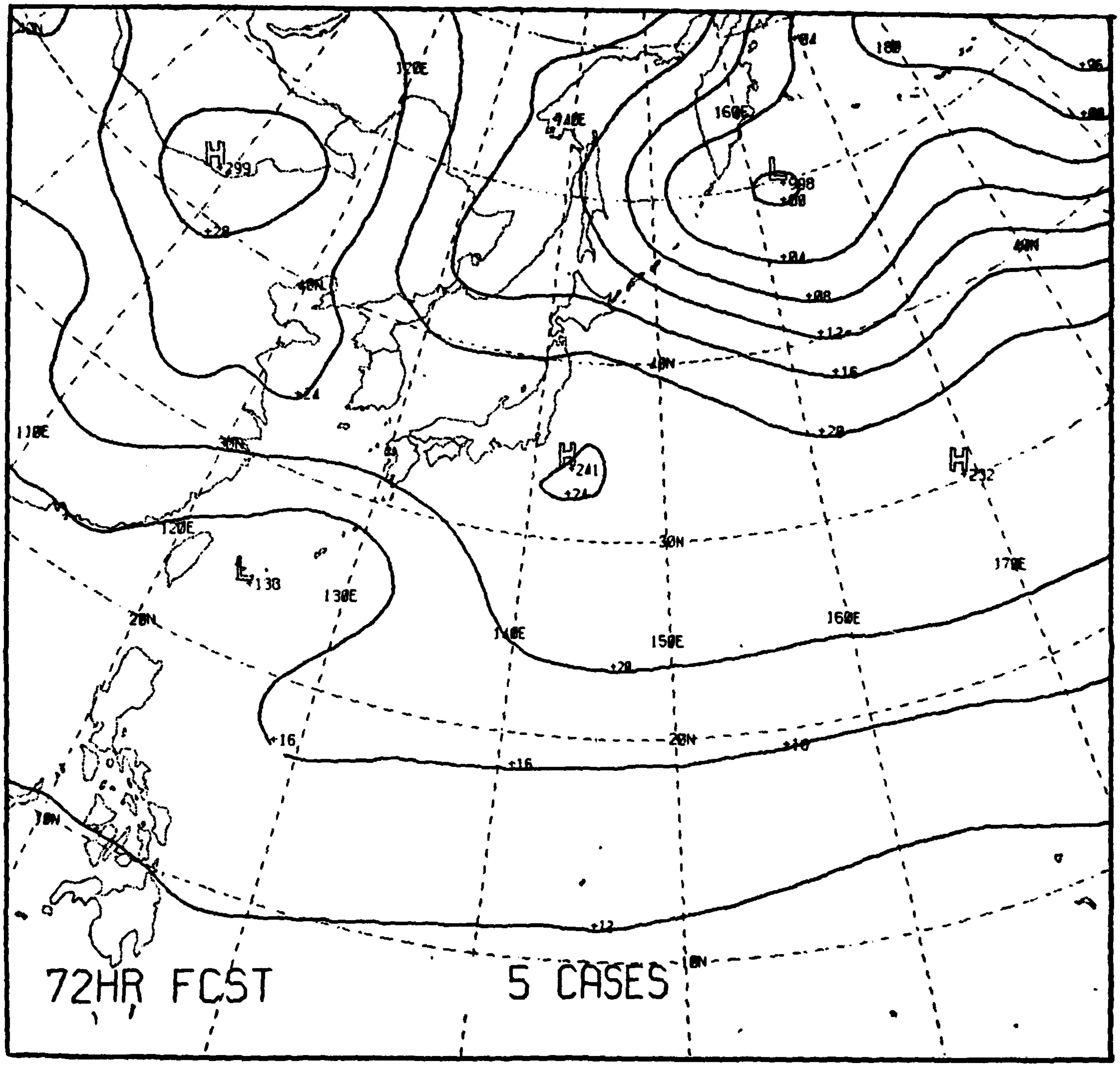


Figure 5f

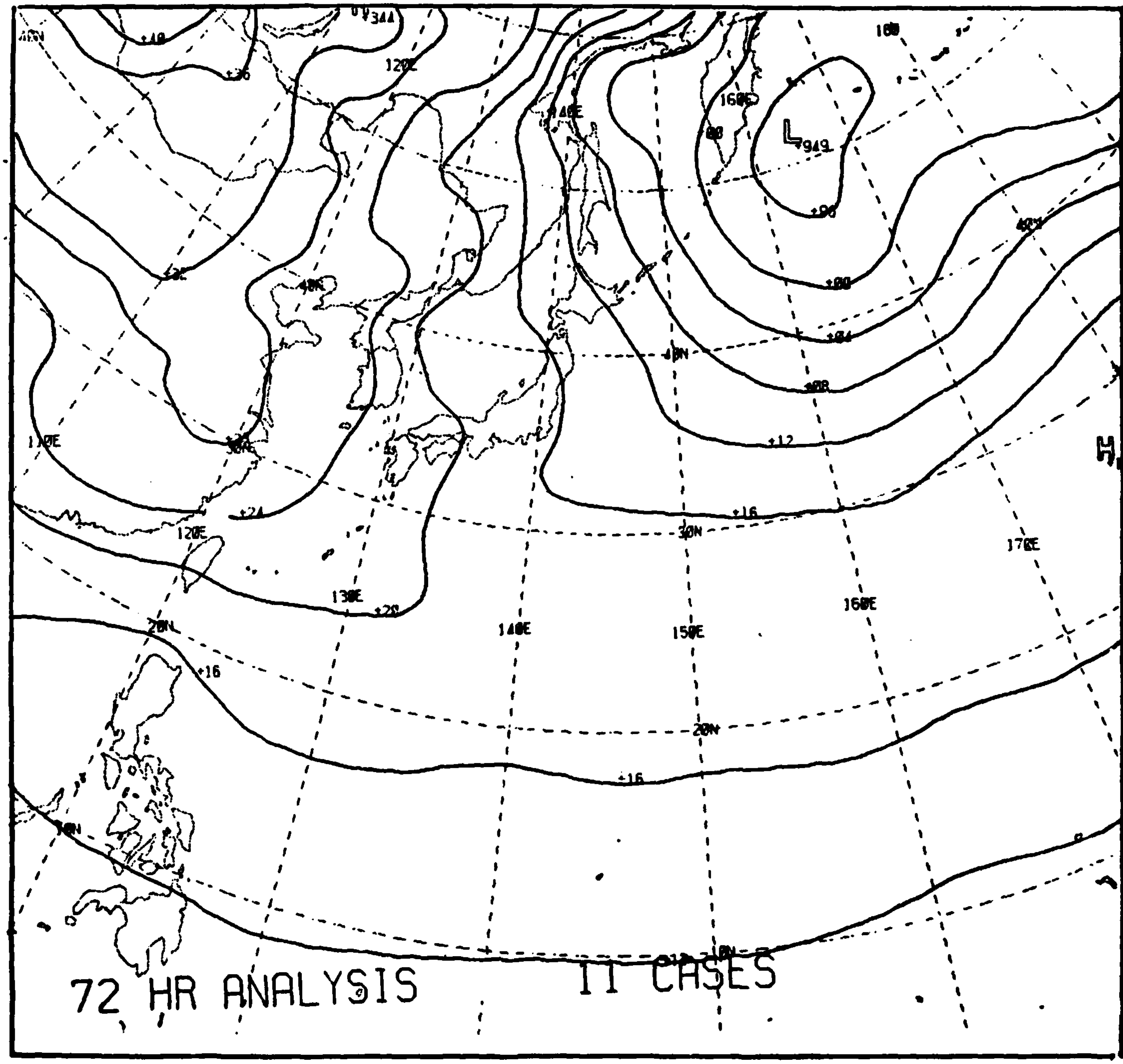


Figure 5g

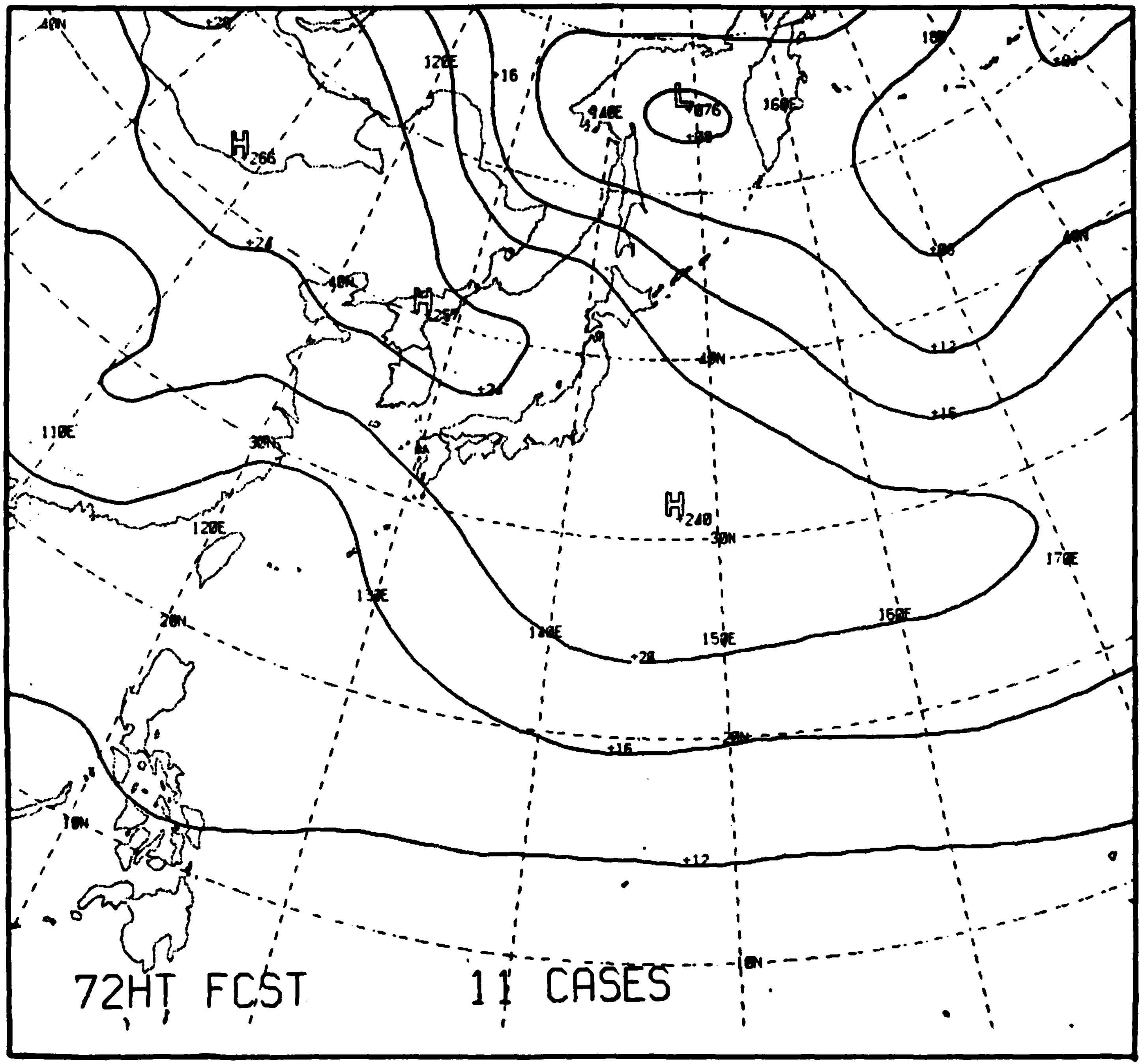


Figure 5h

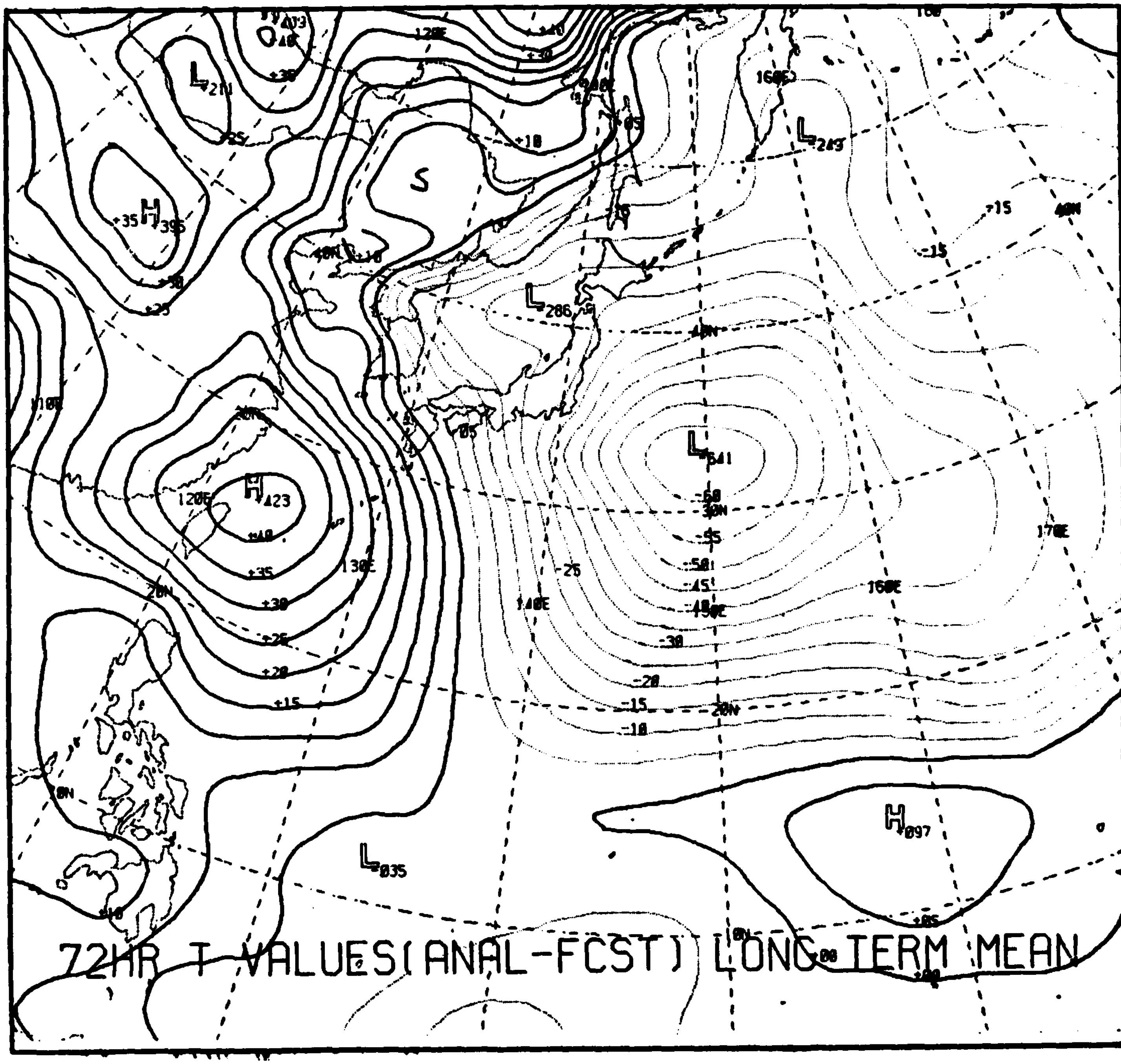


Figure 6 Error pattern constructed of t values for the long term mean group 99% significance level = 2.3

well defined within the region of interest. The error pattern seems to imply the same physical interpretation assigned to the 24 hour pattern. The model is tending to deepen the storms or generally forecast lower pressures than found within the climatological area of formation. The model also tends to under predict the lowering of pressures in the northern area of the region indicating an error in storm movement and development. These are the basic features also found on the 24 hour chart, however the statistical significance is much larger implying an amplification of the error.

Figure 7 shows the error patterns for the remaining short term groups. The 3 case and 4 case groups (figures 7a and 7b) exhibit similar patterns. The patterns are quite noisy and contain many significant error centers. Little similarity is found between these patterns and that of figure 6. The remaining patterns all exhibit a high degree of similarity with the long term mean group of figure 6. The patterns are relatively smooth and contain centers comparable to the mean group.

Figure 8 shows the pattern recognition chart of maximum and minimum centers for the 4, 7 and 11 case groups together with the mean group. This analysis provides the ability to determine which error centers line up with those of the long term group.

The above analysis points out that the 5 case group sufficiently smoothes the smaller scale random errors and provides a relatively smooth systematic error pattern. The randomization of 10 4 case groups was carried out as described in relation to the 24 hour analysis. The randomly determined groups based no resemblance to the longer term groups, thus validating the original 4 case group. This enables the conclusion that an average of 5 cases is sufficient to reduce the random error components of the synoptic prediction system.

These two applications are provided to illustrate the analysis procedure. The same analysis was applied to the

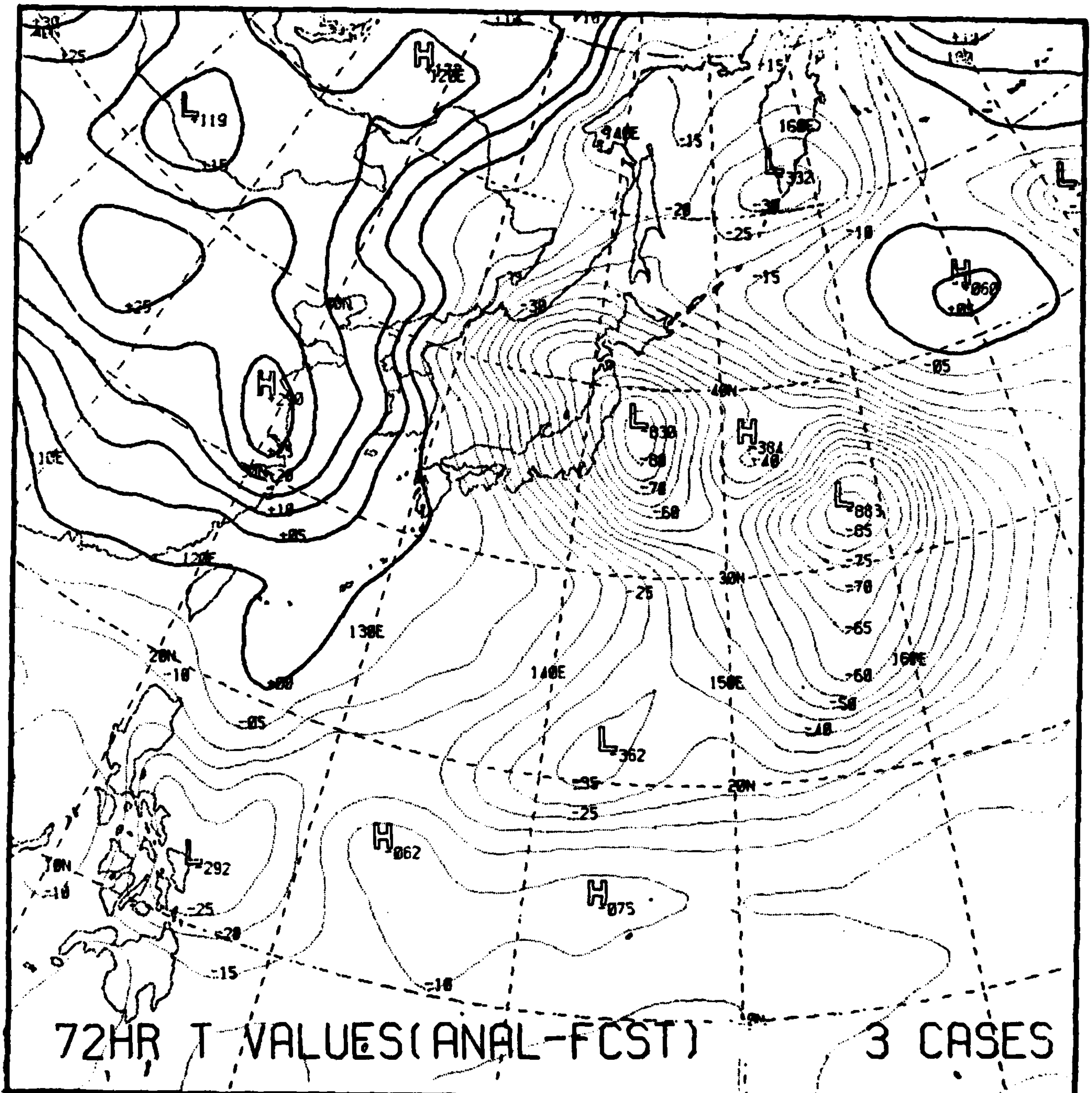


Figure 7 Error patterns for short term groups
Figure 7a 99% significance level = 3.7

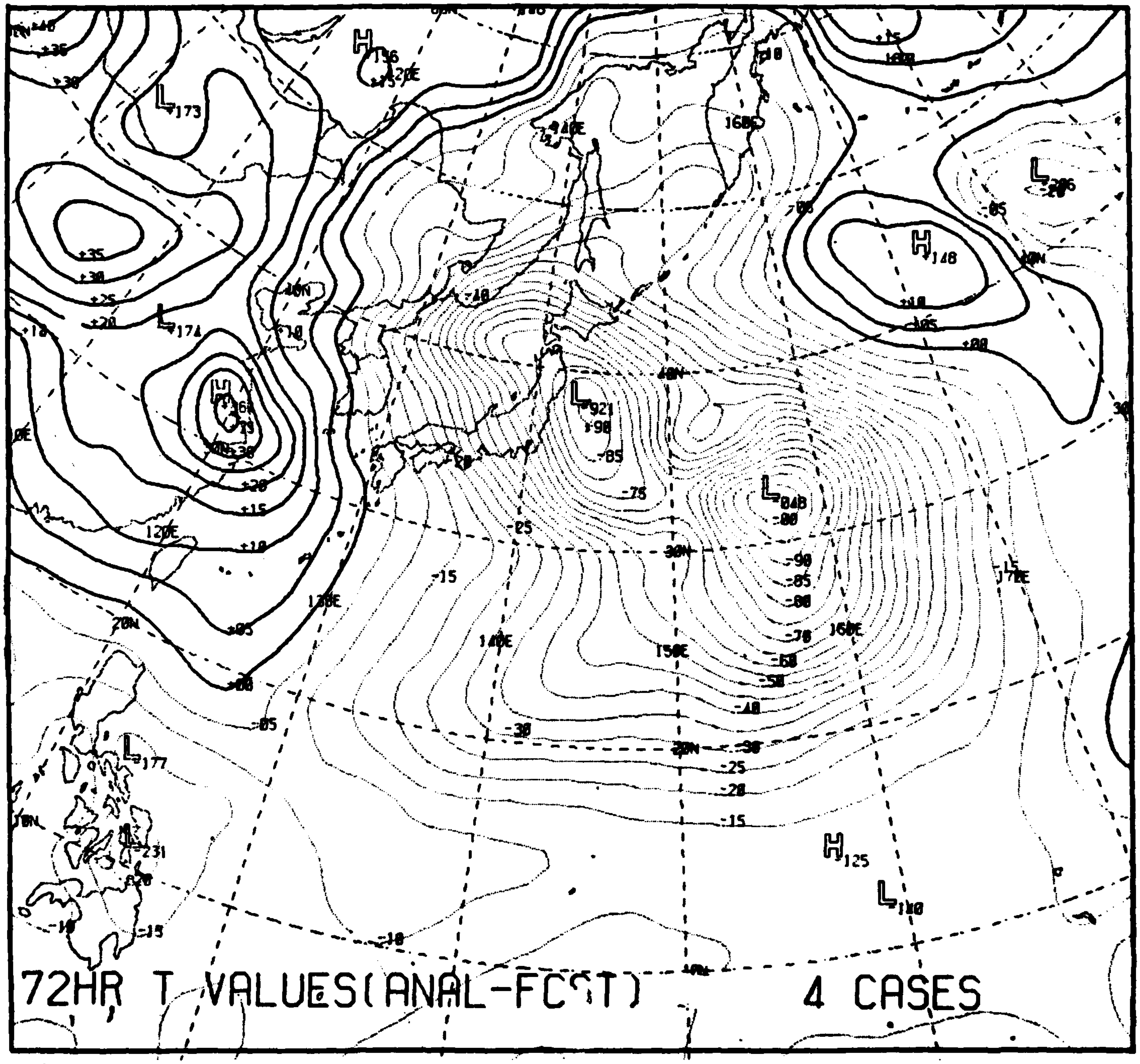


Figure 7b   99% significance level = 3.1

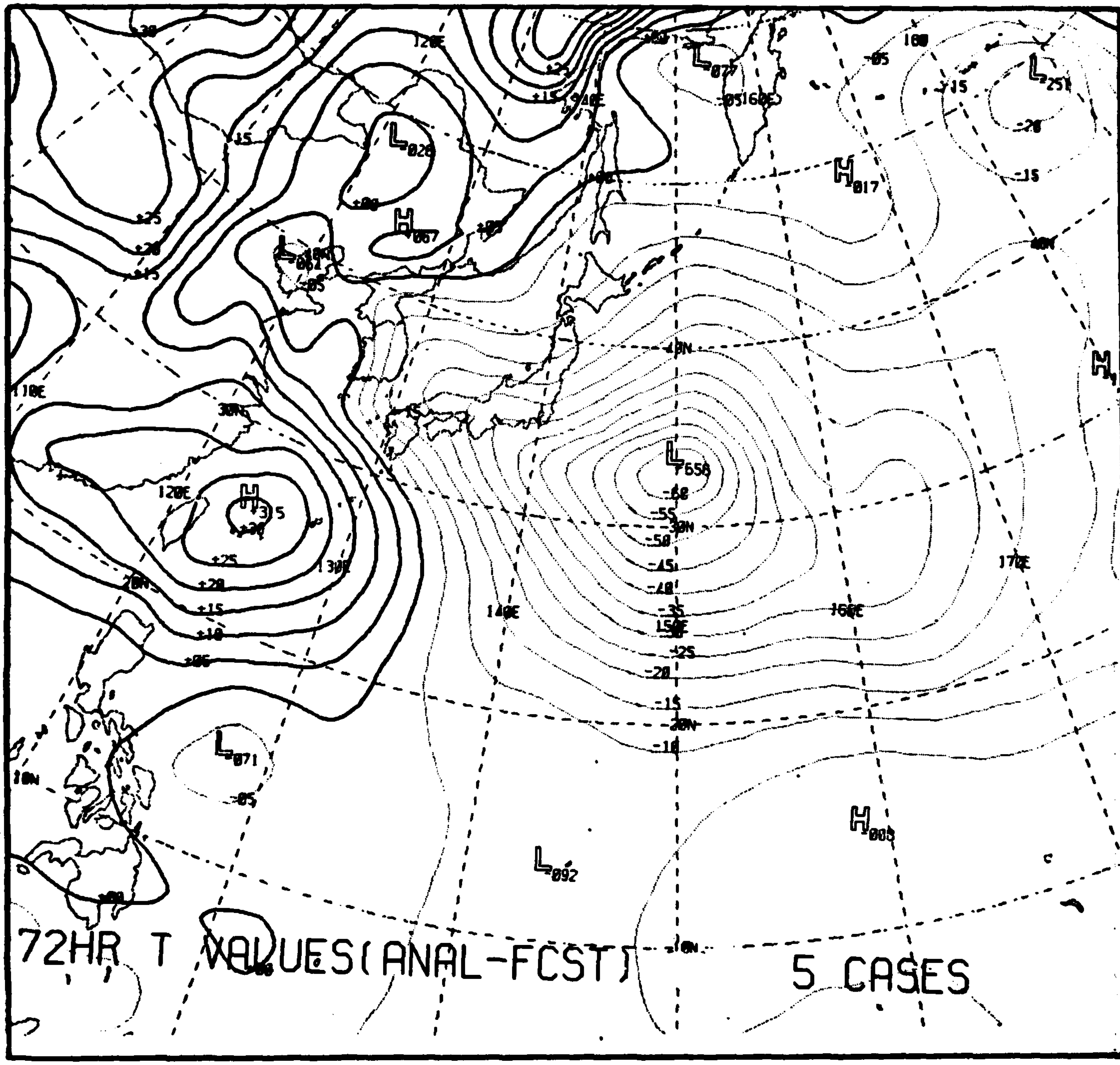


Figure 7c 99% significance level = 2.7

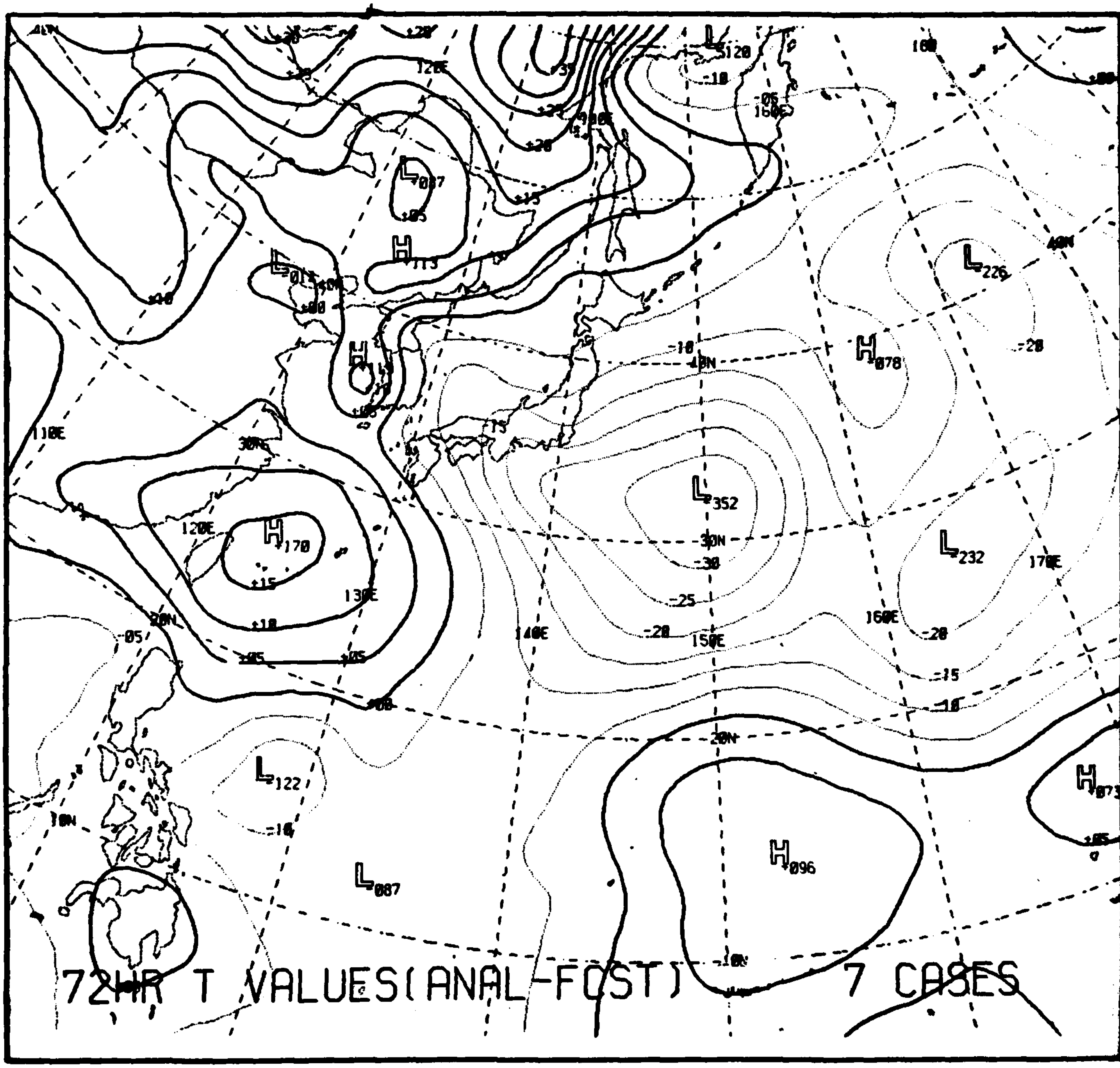


Figure 7d   99% significance level = 2.6

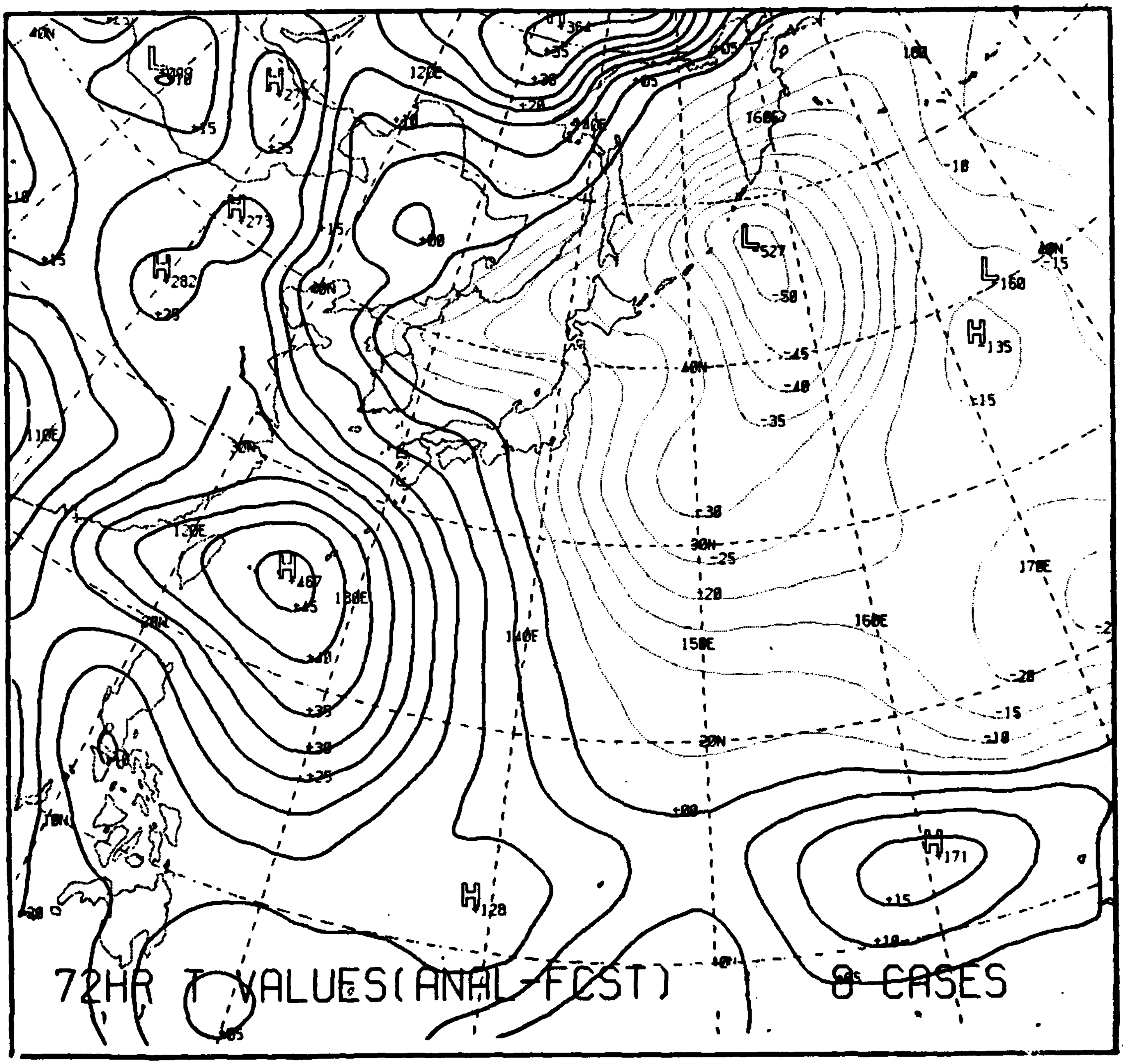


Figure 7d   99% significance level = 2.6

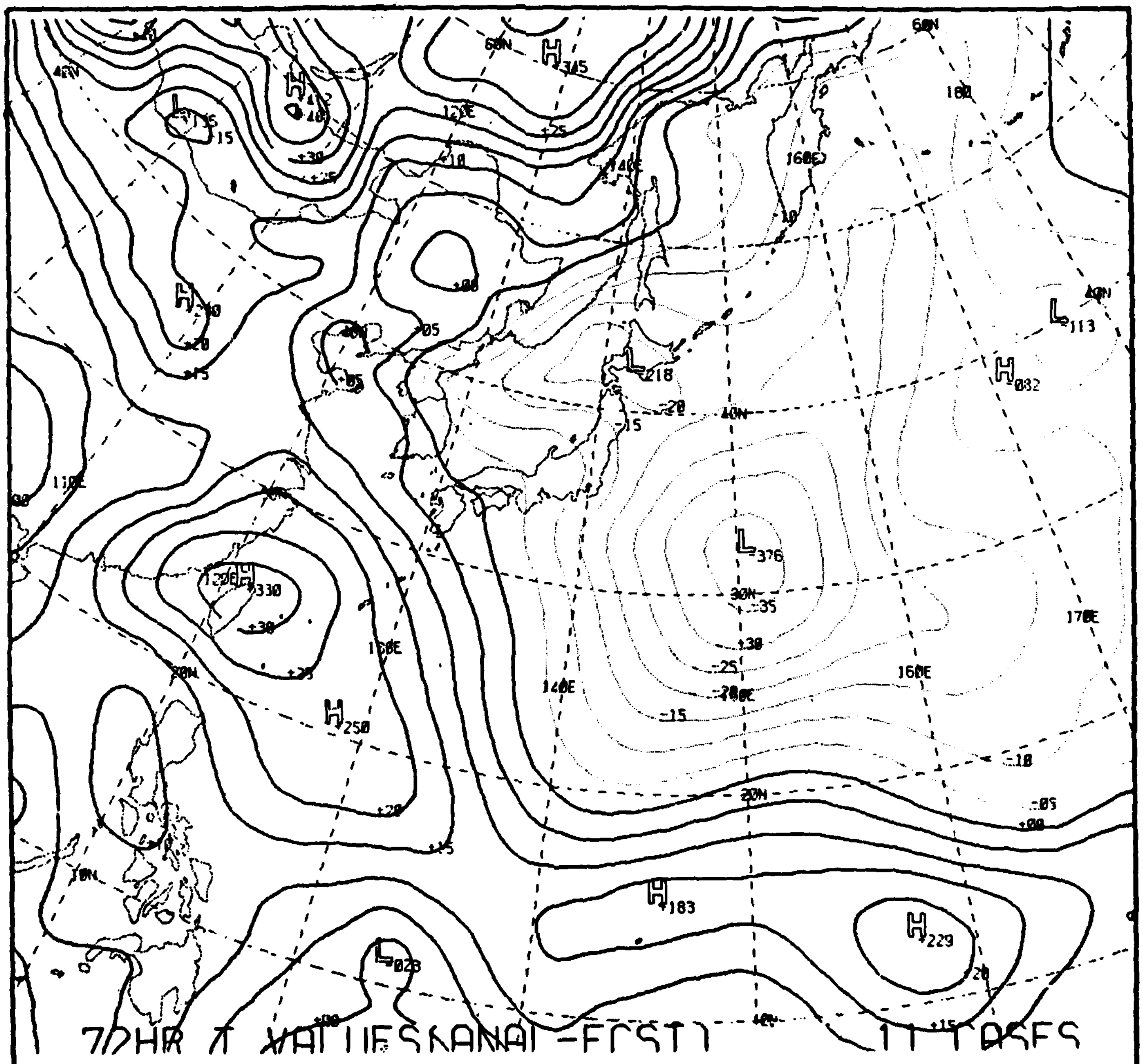


Figure 7f   99% significance level = 2.5

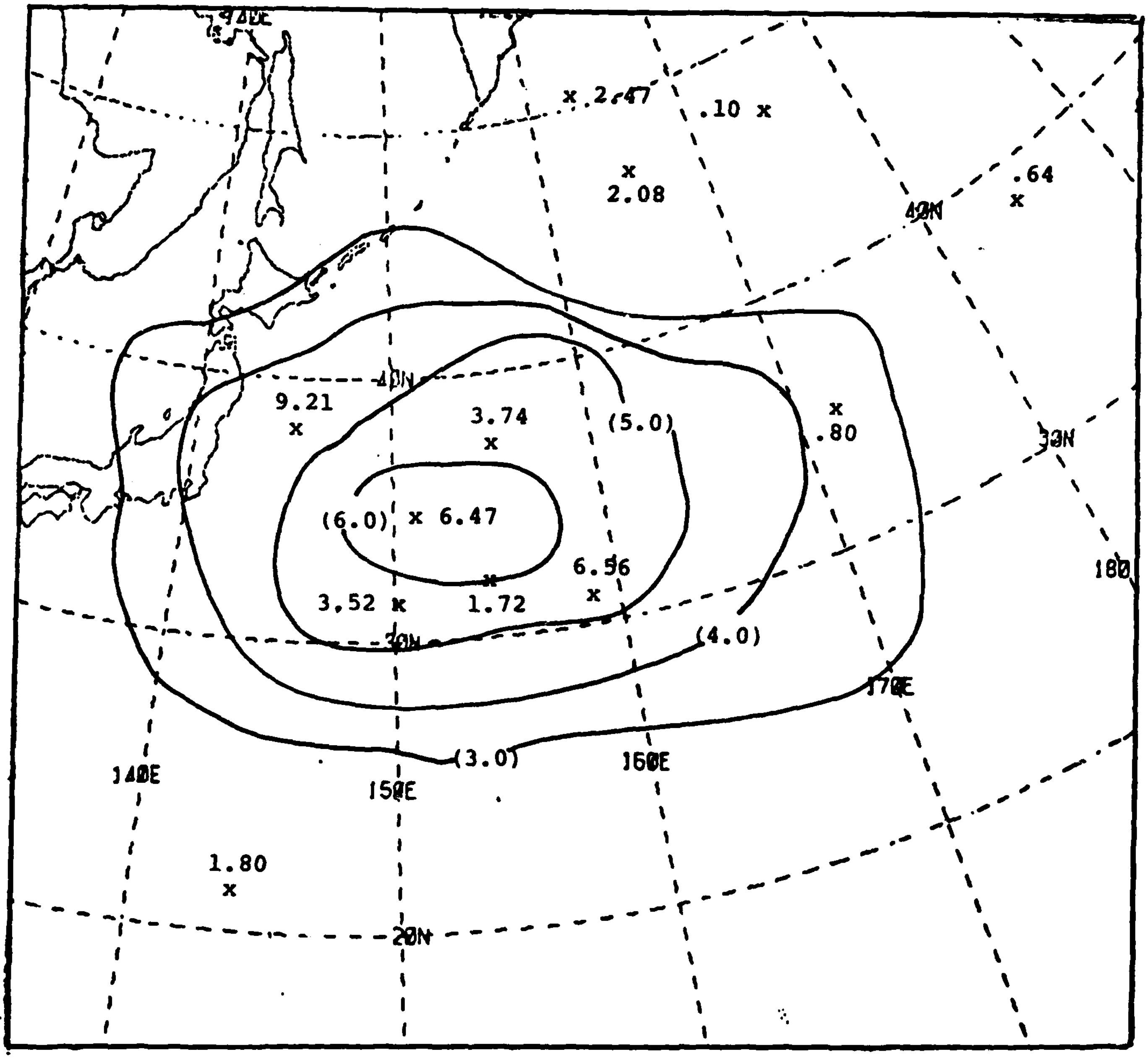


Figure 8 Same as figure 4. Contours represent the 99% significance are of figure 6. Figure 6 error centers are right of the point, figure 7b below the point, figure 7c left of the point and figure 7d below the point

second time periods of the 24 and 72 hour forecasts of the Shanghai Low with similar results. A limited analysis was also applied to the Cape Hatteras Low synoptic event.

4.0 Concluding Comments

This report has documented a statistical procedure which may be used to determine a minimum data period which may be used to average out smaller scale random errors from a field of numerical model weather forecasts. The procedure could be applied for a specific synoptic scale event or sequence of events. For this report, the procedure was applied to a Primitive Equation model's forecast of the Shanghai Low. The procedure would be applied before a more thorough and sophisticated verification analysis or system was to be applied to a numerical forecasting system. This is necessary since many verification methods can not discriminate between systematic or random errors. An application of this methodology may insure that the verification of the numerical model is truly analyzing systematic errors.

One important point which needs to be addressed is the dependency of the results upon the numerical model being used. It cannot be assumed that results obtained for this study would apply to another numerical model. A main advantage of this technique is it is relatively simple and inexpensive to implement. Also it needs to be applied only once.

Results obtained for the application documented here seem reasonable in the statistical and physical sense. The mean error patterns for the 24 and 72 hour forecasts are similar with the 72 hour forecasts more significant. This indicates an amplification of errors as the forecast period is extended.

The minimum data period determined necessary for smoothing out the random errors decreased from the 24 hour to the 72 hour forecasts. This is an indication that the shorter time period forecasts may be generally more noisy in terms of extreme values and gradient.

In summary the described analysis can be used to eliminate random errors from synoptic scale data fields produced from a numerical model forecast. The test applications of the technique has provided reasonable results. The technique could be re-applied to other models before undertaking an extensive verification analysis.

Reference

Somerville, R.C.J., 1977: Pattern recognition for forecast verification. American Meteorological Society Third Conference on Numerical Weather Prediction, Omaha, NE.

Distribution

CHIEF OF NAVAL RESEARCH
LIBRARY SERVICES, CODE 734
RM. 633, BALLSTON TOWER #1
800 QUINCY ST.
ARLINGTON, VA 22217

OFFICE OF NAVAL RESEARCH
CODE 428AT
ARLINGTON, VA 22217

OFFICE OF NAVAL RESEARCH
CODE 420
ARLINGTON, VA 22217

CHIEF OF NAVAL OPERATIONS
OP-952
U.S. NAVAL OBSERVATORY
WASHINGTON, DC 20390

CHIEF OF NAVAL OPERATIONS
NAVY DEPT., OP-986
WASHINGTON, DC 20350

CHIEF OF NAVAL OPERATIONS
ATTN: DR. R. W. JAMES, OP-952D1
U.S. NAVAL OBSERVATORY
34TH & MASS. AVE., NW
WASHINGTON, DC 20390

NAVAL DEPUTY TO THE ADMIN.
NOAA, RM. 200, PAGE BLDG. #1
3300 WHITEHAVEN ST. NW
WASHINGTON, DC 20235

COMMANDING OFFICER
NAVAL RESEARCH LAB
ATTN: LIBRARY, CODE 2620
WASHINGTON, DC 20390

COMMANDING OFFICER
NORDA, CODE 101
NSTL STATION
BAY ST. LOUIS, MS 39529

COMMANDER
NAVOCEANCOM
NSTL STATION
BAY ST. LOUIS, MS 39529

COMMANDING OFFICER
NAVOCEANO, LIBRARY
NSTL STATION
BAY ST. LOUIS, MS 39522

COMMANDING OFFICER
FLENUMOCEANCEN
MONTEREY, CA 93940

COMMANDING OFFICER
NAVWESTOCEANCEN
BOX 113
PEARL HARBOR, HI 96860

COMMANDING OFFICER
NAVEASTOCEANCEN
MCADIE BLDG. (U-117), NAS
NORFOLK, VA 23511

COMMANDING OFFICER
NAVPOLAROCEANCEN
NAVY DEPT.
4301 SUITLAND RD.
WASHINGTON, DC 20390

COMMANDING OFFICER
U.S. NAVOCEANCOMCEN
BOX 12
COMNAVMARIANAS
FPO SAN FRANCISCO 96630

COMMANDING OFFICER
U.S. NAVOCEANCOMCEN
BOX 31
FPO NEW YORK 09540

SUPERINTENDENT
LIBRARY REPORTS
U.S. NAVAL ACADEMY
ANNAPOLIS, MD 21402

CHAIRMAN
OCEANOGRAPHY DEPT.
U.S. NAVAL ACADEMY
ANNAPOLIS, MD 21402

COMMANDER
NAVAIRSYSCOM
ATTN: LIBRARY, AIR-00D4
WASHINGTON, DC 20361

COMMANDER
NAVAIRSYSCOM, AIR-370
WASHINGTON, DC 20361

COMMANDER
PACMISTESTCEN
ATTN: GEOPHYSICS OFFICER, 3250
PT. MUGU, CA 93042

USAFETAC/TS
SCOTT AFB, IL 62225

AFGWC/DAPL
OFFUTT AFB, NE 68113

AFGL/LY
HANSCOM AFB, MA 01731

COMMANDING OFFICER
U.S. ARMY RESEARCH OFFICE
ATTN: GEOPHYSICS DIV.
P.O. BOX 12211
RESEARCH TRIANGLE PARK, NC 27709

DIRECTOR
DEFENSE TECH. INFO. CENTER (DTIC)
CAMERON STATION
ALEXANDRIA, VA 22314

DIRECTOR
SYSTEMS DEVELOPMENT
NWS/NOAA
RM. 1216 - THE GRAMAX BLDG.
8060 13TH STREET
SILVER SPRING, MD 20910

NATIONAL METEOROLOGICAL CENTER
DEVELOPMENT DIVISION
NWS/NOAA
WORLD WEATHER BLDG. W32, RM 204
WASHINGTON, DC 20233

ACQUISITIONS SECTION IRDB-D823
LIBRARY & INFO. SERV. DIV. NOAA
6009 EXECUTIVE BLVD.
ROCKVILLE, MD 20852

NATIONAL WEATHER SERVICE
ATTN: DR. E.W. FRIDAY, DEP. DIR.
GRAMAX BLDG., 8060 13TH ST.
SILVER SPRING, MD 20910

HEAD, ATMOSPHERIC SCIENCES DIV.
NATIONAL SCIENCE FOUNDATION
1800 G. STREET, NW
WASHINGTON, DC 20550

NATIONAL CENTER FOR ATMOS. RSCH.
LIBRARY ACQUISITIONS
P.O. BOX 1470
BOULDER, CO 80302

STATE WATER SURVEY DIVISION
ILLINOIS INST. OF NATURAL
RESOURCES
BOX 5050, STATION A
CHAMPAIGN, IL 61820

EUROPEAN CENTRE FOR MEDIUM RANGE
WEATHER FORECASTS
SHINFIELD PARK, READING
BERKSHIRE RG29AX, ENGLAND